

TITLE	FANTOM Consortium.
JOURNAL	Functional annotation of a full-length mouse cDNA collection
REFERENCE	Nature 409, 685-690 (2001)
AUTHORS	5
TITLE	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
JOURNAL	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
REFERENCE	Nature 420, 563-573 (2002)
AUTHORS	6 (bases 1 to 3126)
	Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanaoka, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Inotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, F., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL: http://genome.gsc.riken.go.jp/ URL: http://fantom.gsc.riken.go.jp/.
FEATURES	Location/Qualifiers
source	1. 3126 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM DB:A730007P14" /db_xref="MGI:2407635" /db_xref="taxon:10090" /clone="A730007P14" /tissue type="cerebellum" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="7 days neonate" 152..1666
CDS	/notes="unnamed protein product; NEURONAL NICOTINIC ACETYLCHOLINE RECEPTOR ALPHA 3 SUBUNIT (SPTR AAL59471, evidence: FASTY, 99.8%ID, 100%length, match=1497) putative" /codon_start=1 /protein_id="BAC37909.1" /db_xref="GI:26348539" /translation="MRSSDMGVVLPPPLMLVLMLLPVSAASEAHLRFQYLPED YNEIIRPVANPEKTLWKPDIVLYNADGDFOVDDKTKALLKYTGVTWIPPAIPKSSCKI QGVFMRVPAEKTLWKPDIVLYNADGDFOVDDKTKALLKYTGVTWIPPAIPKSSCKI DVTFYFDYQNCMTKGSWSYDKAKIDLVLIGSSMLKDYWESGEWAIKAPGKHIEI KYNCEBIYQDITLSLRKPLFYINLIIPCLLSIFLVLFVLPDSCGKVTLCI SVLLSTVFLLVITETIPSLVPLIIGEYLLTFVLSIVITFVLNVHVRTPT HTPTWVAVFLMLPRMFTPTSTEADPKRNFYGAELSNLNCFSRADSKSKE GYPCDGTGCHRRVKIINFSALTRSSSESDVALSLALSPEIKAIQSVKYI AENKQANVAKEIQDDWKYVAMVIDRFLWVFLVCLIGTAGLFLQPLMARDT" 3103..3108 /note="putative" 3126 /note="putative"
	polyA_signal
	polyA_site
ORIGIN	

[illegible]

Db 81 AACGTGTCGACCCAGTCATCATCCATTTTCGAGGTGTCATGTCTCAGCTGTCGAGGTG 140
Qy 220 GAGCAGAGAACAGAGCTTTTAAATAACAAACATCTGGCTAAACATAGAGTGGATGATG 279
Db 141 GATGAAGTAACACAGATCATGGAGACCAACCTGTGGCTCAAGCAAAATCTGGAATGACTAC 200
Qy 280 AACTTGAGTGGAGACACTTCAGATTTCGGCGGGGTCAAGATTTAAGAGTGCACCCAC 339
Db 201 AAGCTGAAGTGGAAACCCCTCGACTATGGTGGGGCAGAGTTCATGCGTGTCCTCGACAG 260
Qy 340 AGACTATGGAACACAGACGTCCTTATGTACAAACAGCGCGACGAAGGGTTTCGACGACG 399
Db 261 AAGATCTGGAAGCAGACATTTGCTGTATAACAAATGCTGTGGGATTTCCAGGTGGAC 320
Qy 400 TATCCACGAAGAGTGGTGGCGGACAAACGGCTCGTGTCTGTACGTGGCGCGCGGATC 459
Db 321 GACAAGACCAAGCCCTTACTCAAGTACACTGGGGAGGTGACTTGGATACCTCCGGCCATC 380
Qy 460 TTCAGAGCACCTGCAAGATCGACATCACCTGGTTCCCTTCGACGACCAACGATGCGAG 519
Db 381 TTTAAGAGCTCTGTAAATCGAGTGCACCTACTTCCCGTTTGATTACCAAAACTGTACC 440
Qy 520 ATGAAGTTTGGCAGCTGGAATTATGATGGTTATCAGTTGGATCTTACAACTACAGGATGAA 579
Db 441 ATGAAGTTTCGTTCTCTAGTACGAGTAAAGCGGAAATCGATCT---GGTCCTGATCGGC 497
Qy 580 GGGGGCGGAGATATAAGCAGTTTGTGCAGAAATGGCGAATGGGAGTTAATAGAGTCCCC 639
Db 498 TCTTCCATGAACCTCAAGGACTATGCGGAGCGCGGAGTGGGCCATCATCAAGGCCCA 557
Qy 640 GGCAAGCGCAACGAGATCTACTCAACTGTGTCGGAGCGCATACATCGACATCAAGTTT 699
Db 558 GGCTACAAACAGACATCAAGTACAACTGTGCGAGAGATCTACCCCGACATCAATAC 617
Qy 700 GCGGTGATGATCGGAGGAAAACGCTCTACTACTTCTTCAATCTGATCGTGCCTGCGTG 759
Db 618 TCGCTGTACATCCGGCGCCTGCTTGTCTACACATCAACCTCATCATCCCTGCGCTG 677
Qy 760 CTCATCGCTCCATGGTCTCTATGGGTTTCACTTGCCTCCAGATCCGAGAAAGTTG 819
Db 678 CTCATCTCTCTCTCACTGTGCTCTTCTACTGCTCCCTCCGACTGCGTGGAGAGTG 737
Qy 820 TCTTTAGGTGTGACGATTAATCTGTGTCGACGGTGTCTCAACATGGTGGCGGAGCG 879
Db 738 ACCCTGTGCAATTTCTGCTCTCTCCCTCCCTGAGCGGTCTCTCTGTTGATCACTGAGACC 797
Qy 880 ATGCCAGCGAGTGGAGCGCGCTTGTGCTCGGCACTTCTTCAACTGCATCATGTTTC 939
Db 798 ATCCCTTCCACCTCGCTGCTCATCCCCCTGATTGGAGAGTACCTCTCTGTTTCAACCATGATT 857
Qy 940 ATGCTGGCTTCTCCGTCGTCCTCCACCATCTGATCTCTCACTTACCAACACCGCGCAGCA 999
Db 858 TTTGTAACTTGTCCATCGTCATCACCGTCTTGTGCTCAACGTGCACTACAGAACCCCG 917
Qy 1000 GACACTCAGAAATGAGTGAATGATTCGTTGGGTGTTCTCTTTATTTGGCTGCGGTGGGTG 1059
Db 918 ACGACACACAAATGCCCTCATGGGTGAAGACTGTATTTCTTGAACCTGCTCCCGAGGTC 977
Qy 1060 CTGGCATGTCAGGCC 1076
Db 978 ATGTTTCATGACGACC 994

RESULT 12
AY402875 1436 bp DNA linear GSS 15-DEC-2003
LOCUS Mus musculus CHRNA3 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY402875
VERSION AY402875.1 GI:39758858
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1436)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Perriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1436)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Perriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..1436
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1436
/gene="CHRNA3"
/locus_tag="HCM1369"
ORIGIN
Query Match 18.1%; Score 269.6; DB 29; Length 1436;
Best Local Similarity 55.1%; Pred. No. 1.1e-41;
Matches 549; Conservative 0; Mismatches 444; Indels 3; Gaps 1;
Qy 81 GGCACGCTGCGGTGATCCACGAGAAGCGGTACTGACACCTATTGGACACCTACACGT 140
Db 2 GGCACGCGCTCGGAAGCTGAGCACCGCCTGTTCCAGTACCTGTTTGAAGATTACAAGA 61
Qy 141 ACTGGAAGCGCCGTCGTCAACGAGAGCGACCCGCTGCGAGCTCTCTTCGCGCTCAGCT 200
Db 62 GATCATCCGCGCGTGGCTAACGTGTGCCATCTCTGTCATCATCCAGTTTGAGGTGTCTAT 121
Qy 201 CATGCAATCATCGAGTGGACGAGAACAGAGCTTTTAAATACAAACATCTGGCTAAA 260
Db 122 GTCTCAGTGGTGAAGGTGGATGAAGTAAACAGATCATGGAACCAACCTGTGGCTGAA 181
Qy 261 ACTAGAGTGAATGATATGAATTTGAGGTGGAAACACTTCAGATTTTCGCGGGGTCAAAGA 320
Db 182 GCAAACTGGAATGACTACAAGCTGAAATGGAAACCTCTGACTTACCAAGGGTGGAGTT 241
Qy 321 TTTAAGAGTGCACCCACAGACTATGGAACCCAGAGCTCTTATGTACAAACGCGCGGA 380
Db 242 CATGCGAGTCCCTGCGAGAGAAGATCTGGAAACAGACATCGTCTTTTACAAACACGCGCA 301
Qy 381 CGAAGGTTTCGACAGCAGCTATCCACGAACGTTGGTGGGAAACCAACGCTCGTCTCT 440
Db 302 TGGGATTTCCAAGTGGATGACAAACCAAGCTCTACTCAAGTACACAGAGAGTGAC 361
Qy 441 GTACGTGCGCCCGCATCTTCAAGAGCACCTGCAAGATCGACATCACTGTTGCCCTT 500
Db 362 TTGGATCCCTCCGGCCATCTTTAAGAGCTCATGCAAAATCGATGTGACCTACTTCCGTT 421
Qy 501 CGACGACCAAGATCGAGATGAAGTTTGGCACTGGACTTATGATGTTTATCAGTTGA 560
Db 422 TGACTACCAAACTGCACCATGAAGTTTCGGCTCCTGCTCTACGACGAAGCAAGATCGA 481
Qy 561 TCTACAACTACAGATGAAGGGCGGAGATATAAGCAGTTTGTCCAGAAATGCGCAATG 620
Db 482 CCT---GGTCCTCATTTGGCTCTCCATGAACCTCAAGGACTATTTGGAAAGTGGCGAGT 538
Qy 621 GGAGTTAATAGGAGTCCCCGGCAAGCGCAACGAGATCTACTACAACTGTTGTCGGAGCC 680

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:09:28 ; Search time 42.6667 Seconds

(without alignments)
3284.615 Million cell updates/sec

Title: US-09-303-232-4

Perfect score: 2665

Sequence: 1 MCGRRARRSHLAAGALLLL.....LFTIIATLAVLLSAPHIMWS 496

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2665	100.0	496	3 AAY50815	Aay50815 H. viresc
2	1850	69.4	770	3 AAY50814	Aay50814 D. melano
3	1803.5	67.7	501	3 AAY50816	Aay50816 H. viresc
4	1417.5	53.2	311	4 ABB63683	Abb63683 Drosophil
5	1267	47.5	502	2 AAW12368	Aaw12368 Neuronal
6	1253	47.0	498	4 ABB60432	Abb60432 Drosophil
7	1246	46.8	502	2 AAW44153	Aaw44153 Human neu
8	1246	46.8	502	2 AAW09025	Aaw09025 Neuronal
9	1246	46.8	502	3 AAB24088	Aab24088 Human PRO
10	1246	46.8	502	4 AAB82690	Aab82690 Nicotinic
11	1246	46.8	502	4 AAB50012	Aab50012 Wild-type
12	1246	46.8	502	5 ABB70492	Abb70492 Human neu
13	1246	46.8	502	5 ABB82435	Abb82435 Human neu
14	1246	46.8	502	7 ADA10874	Ada10874 Human neu
15	1246	46.8	502	7 ADA47051	Ada47051 Human PRO
16	1246	46.8	502	7 ADE57310	Ade57310 Human PRO
17	1242	46.6	502	2 AAW69216	Aaw69216 Mutant hu
18	1240	46.5	502	4 AAB50015	Aab50015 Mutant hu
19	1236	46.4	502	4 AAB50016	Aab50016 Mutant hu
20	1232.5	46.2	502	7 ADE57308	Ade57308 Rat Prote
21	1232.5	46.2	502	7 ADDA7049	Adda7049 Rat Prote
22	1230	46.2	502	4 AAB50017	Aab50017 Mutant hu
23	1219.5	45.8	511	2 AAW12369	Aaw12369 Neuronal
24	1053.5	39.5	554	4 AAE12824	Aae12824 Caenorhab
25	1053.5	39.5	554	6 ABP96318	Abp96318 Caenorhab

ALIGNMENTS

RESULT 1
AAY50815
ID AAY50815 standard; protein; 496 AA.
XX AC AAY50815;
XX AC
DT 17-FEB-2000 (first entry)
XX
DE H. virescens acetyl-choline receptor protein from clone Hva7-1.
XX
DE Acetyl-choline receptor; nicotinic; insect; insecticide; screening;
KW neurotransmission; plant protection agent; conductance; ACHR.
XX
OS Heliothis virescens.
XX
PN DE19819829-A1.
XX
PD 11-NOV-1999.
XX
PF 04-MAY-1998; 98DE-01019829.
XX
PR 04-MAY-1998; 98DE-01019829.
XX
XX (FARB) BAYER AG.
XX PA
XX Adamczewski M, Oellers N, Schulte T;
XX WPI; 2000-014207/02.
DR N-PSDB; AA224476.
XX
XX
PT New nucleic acid encoding a nicotinic acetylcholine receptor from
PT insects, used to identify potential insecticides.
XX
XX Example 1a; Page 17-19; 26pp; German.
XX
XX This invention describes a novel nucleic acid (NA) encoding a nicotinic
XX acetyl-choline receptor (I) from insects which can be used as an
XX insecticide. Inhibitors of (I) interfere with neurotransmission. (I)
XX (also vectors containing it, its regulatory regions, and antibodies
XX directed against (I)-encoded proteins) are used to screen for: (a) plant
XX protection agents that alter conductance of ACHR, potentially useful as
XX insecticides, or (b) genes which encode polypeptides that are involved in
XX formation of functionally related ACHR in insects. (I) are also used to
XX isolate and characterize the specified regulatory regions and for
XX recombinant production of (II). This sequence represents an acetyl-
XX choline receptor isolated from Heliothis virescens
XX Sequence 496 AA;

Aab50014 Chimeric
Aab50018 Mature ce
Abb60716 Drosophil
Aao17243 Modified
Ade17243 Human Pro
Ade57314 Human Pro
Ade57318 Human Pro
Aaw09022 Neuronal
Aad10857 Human neu
Aaw44155 Human neu
Aar73966 Alpha 2 s
Aaw09021 Neuronal
Abg61850 Prostace
Abg31800 Human neu
Abb82430 Human neu
Ada10855 Human neu
Adc71171 Human 205
Adc71169 Human 205
Add45584 Human Pro
Ade59169 Human Pro
Ada83810 Human CHR

Query Match	100.0%;	Score 2665;	DB 3;	Length 496;
Best Local Similarity	100.0%;	Pred. No. 4.5e-257;		
Matches 496;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
1	MGGRARRSHLAAAPAGILLLLCLLWPGRCARGYHEKELHLLHLDHYNVLERPVVNESDPLQ	60		
1	MGGRARRSHLAAAPAGILLLLCLLWPGRCARGYHEKELHLLHLDHYNVLERPVVNESDPLQ	60		
61	LSFGILTLMQIIDDVEKNQLLITNWLKLEWMDNMLRWNTSDFGVKDGLVPPHRLWKPDV	120		
61	LSFGILTLMQIIDDVEKNQLLITNWLKLEWMDNMLRWNTSDFGVKDGLVPPHRLWKPDV	120		
121	LMYNSADEGEFSTYPTTNVVRNNGSCLYVPPGIFKSTCKIDITWPFDDQRCCEMFGSWT	180		
121	LMYNSADEGEFSTYPTTNVVRNNGSCLYVPPGIFKSTCKIDITWPFDDQRCCEMFGSWT	180		
181	YDGVQLDLQLODEGGGDISSFVTNGEWELIGVPGKNEIYINCCPEPYIIDITFVVIRRK	240		
181	YDGVQLDLQLODEGGGDISSFVTNGEWELIGVPGKNEIYINCCPEPYIIDITFVVIRRK	240		
241	TLXYFFNLIIVPCVLIIASWALLGFTLPPDSGEKLSLGVTTILSLTVELNVAETMPATSDA	300		
241	TLXYFFNLIIVPCVLIIASWALLGFTLPPDSGEKLSLGVTTILSLTVELNVAETMPATSDA	300		
301	VPLLGTYNCFIMFVASSVSTIILINYNHRRHADTHEMSDWIRCVFLYWL PWVLRMSRPG	360		
301	VPLLGTYNCFIMFVASSVSTIILINYNHRRHADTHEMSDWIRCVFLYWL PWVLRMSRPG	360		
361	SATTPPPARVPPDPDLRLRERSKSLIANVLIDDDDFRHPQAQQPQCCRYRYGGENGAG	420		
361	SATTPPPARVPPDPDLRLRERSKSLIANVLIDDDDFRHPQAQQPQCCRYRYGGENGAG	420		
421	LAASCFGVDYELSLIIKEIRVITDQMKODEDADISRWKFAAMVVDRLCLIIIFTLFTI	480		
421	LAASCFGVDYELSLIIKEIRVITDQMKODEDADISRWKFAAMVVDRLCLIIIFTLFTI	480		
481	IATLAVLISAPHIMVS	496		
481	IATLAVLISAPHIMVS	496		

PT insects, used to identify potential insecticides.

XX PS Example 1a; Page 12-14; 26pp; German.

XX This invention describes a novel nucleic acid (NA) encoding a nicotinic
CC acetyl-choline receptor (I) from insects which can be used as an
CC insecticide. Inhibitors of (I) interfere with neurotransmission. (I)
CC (also vectors containing it, its regulatory regions, and antibodies
CC directed against (I)-encoded proteins) are used to screen for: (a) plant
CC protection agents that alter conductance of AChR, potentially useful as
CC insecticides, or (b) genes which encode polypeptides that are involved in
CC formation of functionally related AChR in insects. (I) are also used to
CC isolate and characterize the specified regulatory regions and for
CC recombinant production of (II). This sequence represents an acetyl-
CC choline receptor isolated from *Drosophila melanogaster*

QY 242 LYFFNLIVPCVLIASMLLQFTLPDPDSGEKLSGLVTILLSLTVFLNVAETMPATSDAV 301
 DB 247 LYFFNLIVPCVLIASMLLQFTLPDPDSGEKLSGLVTILLSLTVFLNVAETMPATSDAV 306
 QY 302 PLLG 305
 DB 307 PLLG 310
 RESULT 5
 AAW12368
 ID AAW12368 standard; protein; 502 AA.
 AC AAW12368;
 XX
 DT 17-JUN-1997 (first entry)
 DE Neuronal alpha-bungarotoxin binding protein alpha 1; cholinergic;
 KW ligand binding; ion channel.
 XX Gallus sp.
 OS
 FH Key Location/Qualifiers
 FT Peptide 1..22
 FT Protein /label= Sig_peptide
 FT 23..502
 FT /label= Mat_protein
 XX
 PN US5599709-A.
 XX
 PD 04-FEB-1997.
 XX
 PF 28-SEP-1989; 89US-00413947.
 XX
 PR 28-SEP-1989; 89US-00413947.
 XX
 PA (SALK) SALK INST BIOLOGICAL STUDIES.
 XX
 PI Lindstrom JM, Schoepfer RD;
 XX
 DR WPI; 1997-118297/11.
 DR N-PSDB; AAT59196.
 XX
 PT New isolated neuronal alpha-bungarotoxin-binding protein DNA - used to
 screen cholinergic agents and other drugs which may affect ligand
 binding, ion channel or other activities of the protein.
 XX
 PS Example; Fig 2A-B; 18pp; English.
 XX
 CC The alpha subunit (AAW12368) and alpha2 subunit (AAW12369) amino acid
 sequences of chick neuronal alpha-bungarotoxin binding protein (ABBP)
 were deduced from newly isolated DNA molecules (AAT59196-97) obtd. from
 an 18-day chick embryo cDNA library. ABBP subunits can be produced in
 recombinant host cells, pref. a bacterium, and used in the screening of
 cholinergic agents and other drugs that may affect the ligand binding,
 ion channel or other activity of intact ABBP subtypes. The ABBP alpha1
 and alpha2 subunits can also be used to produce subunit peptides for use
 as immunogens for preparing antibodies to permit affinity purification of
 subtypes and their histological location
 XX
 SQ Sequence 502 AA;
 Query March 47.5%; Score 1267; DB 2; Length 502;
 Best Local Similarity 48.6%; Pred. No. 2.5e-117;
 Matches 249; Conservative 83; Mismatches 144; Indels 36; Gaps 6;
 QY 1 MGRARRSLAAPAGLLLLCLLWPRGARGCYHEKRLHLHLHDYVNLVRPVNVEDPLQ 60
 DB 1 MGLRALMLWLLAAAGLV-----RESLQGEFQRLKYLKYNPLRPFVANDSQPLT 52
 QY 61 LSFGLTLMQIIDVDEKNQLITNIWLKLEWMDMLNWNWTSDFGGVKDLRVPFHLRWKPDV 120

DB 53 VYFTLSLMQIMDVDEKNQVLTTNIWLQMYWTDHYLQWNVSEYPCGVKNVRFDPDGLIWKPI 112
 QY 121 LMVNSADEGSDSTPTNVVVRNNGSCLYVPPGIFKSTCKIDITWFPDDQRCENMKFGSWT 180
 DB 113 LLVNSADERFDATEHTNLVNSSGHCQYLPPIGFKSSCYIDVRWFPFDVQKCNLKFSGWT 172
 QY 181 YDGYQLDLQLODEGGDISFVTNGEWELIGVPGKRNEIYVNCPCPEPIDITFAVWIRK 240
 DB 173 YGWSLDLQGE---ADISGISGEWDLVGIPKRTESFYECCKEYPDITFTVWRRR 229
 QY 241 TLYYFFNLIVPCVLIASMLLQFTLPDPDSGEKLSGLVTILLSLTVFLNVAETMPATSDA 300
 DB 230 TLYYGLNLLIPCVLISALALLVFLLPADSGEKISLIGITVLLSLTVFLNVAEIWPAISDS 289
 QY 301 VPLLGTFFNCIMFVWASSVSTILINHYHRHATHEDMSDWIRCVFLVWLVLRMSRPG 360
 DB 290 VPLIAQYFASTMIIVGLSVVTVIQLQHHHDPDGGKMPKWTVRVILLNWCWFLRMKRP 349
 QY 361 SATTPPPARVPPPPDLELRSSKSLLANVLID-----IDDFRHPQAQQ 404
 DB 350 E-----DKVRPACQHKQRCSSLSSMEMNTVSGQCSNGNMLYIGFRCLDGVHCTPTTDS 403
 QY 405 PQCRYRGG--EENGAGLAHSCFVDYELSLILKEIRVITDQMRKXDEADISRDWKF 462
 DB 404 GVICGRMTCSPTEENLLHSGHPSEG-DPDLAKILEEVRYIANRFRDQDEBEAICNEWK 462
 QY 463 AAMVVDRLCLLIIFTLTIATLAVLSAPHIM 494
 DB 463 AASVVDRLCLMAFSVFTIICTIGILMSAPNFV 494
 RESULT 6
 ABB60432
 ID ABB60432 standard; protein; 498 AA.
 XX
 AC ABB60432;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 8088.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL04535.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 genes from Drosophila and for elucidating cell signalling and cell-cell
 interactions.
 XX
 PS Disclosure; SEQ ID NO 8088; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 capable of detecting 1000 or more genes from Drosophila. The invention is
 useful in developmental biology and in elucidating cell signalling and
 cell-cell interactions in higher eukaryotes for the development of

[illegible]

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 8, 2004, 06:45:49 ; Search time 5459.09 Seconds
(without alignments)
3938.047 Million cell updates/sec

Title: US-09-303-232-4

Perfect score: 2665

Sequence: 1 MGRARRSHAPAGLLLLL.....LFTIATLAVLLSAPHIMVS 496

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09303232/runat_07052004_101110_23883/app.query.fasta_1.2261
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAFSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09303232 @CGN 1.1 8225 @runat_07052004_101110_23883 -NCFU=6 -ICFU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.em.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg_hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htg_mus.*
34: em.htg_pln.*
35: em.htg_rod.*
36: em.htg_man.*
37: em.htg_vrt.*
38: em.sy.*
39: em.htgo_hum.*
40: em.htgo_mus.*
41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2665	100.0	3629	3	AF143846 Heliothis
2	2665	100.0	3700	6	AX009612 Sequence
3	2665	100.0	3701	6	E58347 Nuclaic aci
4	2064.5	77.5	2907	3	AF272778 Drosophil
5	2047	76.8	1683	3	DMES54210
6	1912	71.7	2834	3	AY036613 Drosophil
7	1850	69.4	2886	6	E58346 Nuclaic aci
8	1850	69.4	2886	6	AX009610 Sequence
9	1809	67.9	3029	3	AF143847 Heliothis
10	1809	67.9	3109	6	E58348 Nuclaic aci
11	1809	67.9	3109	6	AX009614 Sequence
12	1795.5	67.4	2023	3	AF21446 Drosophil
13	1792.5	67.3	2023	3	AF321447 Drosophil
14	1791.5	67.2	1665	3	DMES54209
15	1791.5	67.2	1699	3	BT011147 Drosophil
16	1790.5	67.2	2023	3	AF321445 Drosophil
17	1774	66.6	2068	3	AF321448 Drosophil
18	1769	66.4	2110	3	AF321449 Drosophil
19	1386.5	52.0	2116	3	AY036614 Drosophil
20	1274.5	47.8	2037	5	GNNARA7A
21	1272.5	47.7	1719	5	AY247962 Danio rer
22	1267	47.5	2696	5	GGANAREC
23	1253	47.0	1668	5	AY298752 Takifugu
24	1250	46.9	1848	10	MUSNARS
25	1248	46.8	2088	10	AF225980
26	1247	46.8	1876	6	AR055255 Sequence
27	1247	46.8	1876	6	AR071403 Sequence
28	1247	46.8	1876	6	AR173187 Sequence
29	1247	46.8	1876	6	AR224030 Sequence
30	1247	46.8	1876	6	AR282833 Sequence
31	1247	46.8	1876	6	AX719088 Sequence
32	1247	46.8	1876	9	HSU62436 Human nicot
33	1246	46.8	1509	6	AX054567 Sequence
34	1246	46.8	1509	9	HSNACHRA7
35	1242	46.6	1590	6	BD021850 Sequence
36	1242	46.6	1590	6	BD023656 Variant h
37	1242	46.6	1916	9	AF385585 Homo sapi
38	1242	46.6	1977	9	AF385585 Homo sapi
39	1242	46.6	2106	10	RATNARAD
40	1241	46.6	1559	9	HUMAYNAR
41	1241	46.6	2087	9	HSARA7A
42	1240	46.5	1509	6	AX054575 Sequence
43	1239.5	46.5	1461	5	AY299465 Takifugu
44	1237.5	46.4	2107	9	AF486623 Macaca mu
45	1237.5	46.4	3030	10	S53987 nicotinic r

ALIGNMENTS

RESULT 1

AF143846 3629 bp mRNA linear INV 27-MAY-1999
 LOCUS Heliothis virescens putative nicotinic acetylcholine receptor alpha
 DEFINITION 7-1 subunit mRNA, complete cds.
 ACCESSION AF143846
 VERSION AF143846.1 GI:4895004
 KEYWORDS SOURCE
 ORGANISM Heliothis virescens (tobacco budworm)
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 Noctuoidea; Noctuidae; Heliothinae; Heliothis.
 REFERENCE 1 (Bases 1 to 3629)
 AUTHORS Schulte, T., Oellers, N. and Adamczewski, M.
 TITLE Putative alpha subunits of insect nicotinic acetylcholine receptors
 more similar to vertebrate alpha 7 subunits and C. elegans Ce21
 than to other insect nicotinic acetylcholine receptor alpha
 subunits
 JOURNAL Unpublished
 REFERENCE 2 (Bases 1 to 3629)
 AUTHORS Schulte, T., Oellers, N. and Adamczewski, M.
 TITLE Direct Submission
 JOURNAL Submitted (19-APR-1999) ZF-BTB, Bayer AG, Bldg. Q 18,, Leverkusen
 51368, Germany
 FEATURES Location/Qualifiers
 1..3629
 /organism="Heliothis virescens"
 /mol_type="mRNA"
 /db_xref="taxon:7102"
 335..1825
 /note="hvnachra-7-1"
 /codon_start=1
 /product="putative nicotinic acetylcholine receptor alpha
 7-1 subunit"
 /protein_id="AAD32697.1"
 /db_xref="GI:4895005"
 /translation="MGRARRSLAAPAGLLLLCLLWPRGARGCYHEKRLHLLDH
 YNLEFVNESDPQLSFGILTMQIIDVEKNQLLITNIWLKLEWDMNLWNSTDF
 GVKVLDVPHRLWKPVLMVNSADEGFDSTPTNVVVRNNGSLYVPPGIFKSTKI
 DIWFFPDORCKBMKGSWTVDGVLQLOLQDEGGDISSPVTNGEWELIGVPGKNE
 IYNCCPEPIDITFAVIRKTLVYFNLVPCVLASMALLOFTLPPSGEXKLSLG
 VTLLSLTVFLNVAETMPATSDAVPLGLTFNCIMFVWASSVSTILILNIHRRHAD
 THEMDIRCVFLVFLSWLRMSRPGSATPPPARVPPPDLELRERSSKSLANVLD
 IDDFRFPQAPQCCRYRGENGAGLAHSCFVDYELSLILKEIRVITDQMRKD
 DEDADISRWKFAAMVVDRLIIFLTFTIATLAVLLSAPHIMVS"
 ORIGIN
 Alignment Scores:
 Pred. No.: 5.82e-247 Length: 3629
 Score: 2665.00 Matches: 496
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0
 US-09-303-232-4 (1-496) x AF143846 (1-3629)
 QY 1 MetGlyGlyArgAlaArgSerHisLeuAlaAlaProAlaGlyLeuLeuLeuLeu 20
 Db 335 ATGGCGCGCGCGCGCGCGCTCGCACTTGGCGCGCGCGCGCGCTGCTGCTGCTG 394
 QY 21 CysLeuLeuTrpProArgGlyAlaArgCysGlyTyrHisGluLysArgLeuHisHis 40
 Db 395 TGCCTGCTTGGCGCGCGCGCGCGCTCGCGGTACCACGAGAGCGGCTACTGCAC 454
 QY 41 LeuLeuAspHisTyrAsnValLeuGluArgProValValAsnGluSerAspProLeuGln 60
 Db 455 CTATTGACCACTACACAGTACTCGAGAGCGCGCTGCTCAACGAGACGCCCGCTCAG 514
 QY 61 LeuSerPheGlyLeuThrLeuMetGlnIleAlaAspValAlaAspGluLysAsnGlnLeu 80
 Db 515 CTCTCTTGGCGCGCTCACGCTCATGCAGATCATCGACGTGGACGAGAGACCGCTTTA 574

QY 81 IleThrAsnIleTyrLeuLysLeuGluTrpAsnAspMetAsnLeuArgTyrAsnThrSer 100
 Db 575 ATACAAACATCTGGCTAAACCTAGAGTGGATGATGATGAATGAATGAGGTGGACACTTCA 634
 QY 101 AspPheGlyGlyValLysAspLeuArgValProProHisArgLeuTyrLysProAspVal 120
 Db 635 GATTTGGCGGGGTCAAAGATTAAAGAGTGCACCCACAGACTATGGAACACGACGCTC 694
 QY 121 LeuMetTyrAsnSerAlaAspGluGlyPheAspSerThrTyrProThrAsnValValVal 140
 Db 695 CTTATGTACAAACAGCGCGGACGAAGGTTTCACACGACGATTCACCAACGACGTTGGTGT 754
 QY 141 ArgAsnAsnGlySerCysLeuTyrValProProGlyLysPheLysSerThrCysLysIle 160
 Db 755 CGGAACAAACGCTGCTGCTGTACGTGCCGCCCGGCATCTTCAAGACGACCTGCAAGATC 814
 QY 161 AspIleThrTyrPheProPheAspAspGlnArgCysGluMetLysPheGlySerTyrThr 180
 Db 815 GACATCACTGGTTCCCTTCGACGACCAACGATCGAGATGAAGTTTGGCAGCTGGACT 874
 QY 181 TyrAspGlyTyrGlnLeuAspLeuGlnLeuAspGluGlyGlyAspIleSerSer 200
 Db 875 TATGATGTTATCATGTTGGATCTCAACTACAGGATGAAGGGGGCGGAGATATAGCAGT 934
 QY 201 PheValThrAsnGlyGluTrpGluLeuIleGlyValProGlyLysArgAsnGluIleTyr 220
 Db 935 TTGTCTACGAATGGCGAATGGGAGTTAATAGGAGTCCCGGCAAGCGCAACGAGATCTAC 994
 QY 221 TyrAsnCysCysProGluProTyrIleAspIleThrPheAlaValIleArgArgLys 240
 Db 995 TACAACCTGTTGTCGGGACCATACATCGACATCATCGTTGCGGTGGTGTATCCGAGGAAA 1054
 QY 241 ThrLeuTyrTyrPhePheAsnLeuIleValProCysValLeuIleAlaSerMetAlaLeu 260
 Db 1055 ACGCTCTACTACTTCTTCAATCTGATCGTGCCTTCGCTGCTCATCGCTCCATGGCTCTA 1114
 QY 261 LeuGlyPheThrLeuProAspSerGlyGluLysLeuSerLeuGlyValThrIleLeu 280
 Db 1115 TTGGGGTTACCTTGGCTCCAGACTCCGAGAGAAAAGTTGCTTTAGGTGTGACCATATTA 1174
 QY 281 LeuSerLeuThrValPheLeuAsnMetValAlaGluThrMetProAlaThrSerAspAla 300
 Db 1175 CTGTCGTGTAGCGGTTCCTCAACATGTCGGCGAGACGATGCCAGCGACGCTCGACGCC 1234
 QY 301 ValProLeuLeuGlyThrTyrPheAsnCysIleMetPheMetValAlaSerSerValVal 320
 Db 1235 GTGCCCTTGTTCGGCACCTCTTCACTGATCACTGATTCATGGTGGCTTCTCCGCTGCTC 1294
 QY 321 SerThrIleLeuIleLeuAsnTyrHisHisArgHisAlaAspThrHisGluMetSerAsp 340
 Db 1295 TCCACCATACTGATCTCAACTACCAACCGGCGACGACGACACTCAGAAATGAGTGAT 1354
 QY 341 TrpIleArgCysValPheLeuTyrTrpLeuProTrpValLeuArgMetSerArgProGly 360
 Db 1355 TGGATTCGTTCGCTGTTCTCTTTATTGCTTCGCTGGGTGCTGCGCATGTTCACGCCCGC 1414
 QY 361 SerAlaThrThrProProAlaArgValProProProProAspLeuGluLeuArgGlu 380
 Db 1415 TCGCGGACGACGCGCGCGCGCGGTACTCCCGCGCGGACCTCGAGACTCGCGGAG 1474
 QY 381 ArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAspAspPheArgHisPro 400
 Db 1475 CGCTCTCTCAAGTCGCTCTTAGCGAACGTCGTCGACATCGATGACGACTTCCGCGCACCC 1534
 QY 401 GluAlaGlnGlnProGlnCysCysArgTyrTyrArgGlyGlyGluGluAsnGlyValGly 420
 Db 1535 CAAGCGCAGCAGCGCAATCTCCGATACTACAGGGGGGGGTGAGGAGAAATGGCGCGGG 1594
 QY 421 LeuAlaAlaHisSerCysPheGlyValAspTyrGluLeuSerLeuIleLeuLysGluIle 440
 Db 1595 TTGGCGCGCGCAGTGTGCTTCGGTGTGCACTACGAGCTCTCCCTCATCTTCGAGAGATT 1654
 QY 441 ArgValIleThrAspGlnMetArgLysAspGluAspAlaAspIleSerArgAspTrp 460

```
Db      1655 AGAGTCATCACAGATCAGATGCGCAGGACGAGAGATTCGCGCGACTGG 1714
Qy      461  LysPheAlaAlaMetValValAspArgLeuCysLeuIlePheThrLeuPheThrIle 480
Db      1715 AAGTTCGCCGCCATGTCGTGGACAGACTGTGCTTATTATCTTACCCTGTTTCAACAATC 1774
Qy      481  IleAlaThrLeuAlaValLeuLeuSerAlaProHisIleMetValSer 496
Db      1775 ATCGCACCGTAGCCGTGCTGTGTCGCGCCACACATCATGCTGTGTCG 1822

RESULT 2
AX009612
LOCUS      AX009612          3700 bp      DNA      linear      PAT 06-SEP-2000
DEFINITION Sequence 3 from Patent EP0962528.
ACCESSION  AX009612
VERSION     AX009612.1  GI:9996844
KEYWORDS   Heliothis virescens (tobacco budworm)
SOURCE      Heliothis virescens
ORGANISM    Heliothis virescens
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
            Noctuoidea; Noctuidae; Heliothinae; Heliothis.
REFERENCE   1
            Adamczewski, M.D., Schulte, T.D. and Oellers, N.D.
            Nucleic acids encoding acetylcholin-receptor subunits from insects
            Patent: EP 0962528-A 3 08-DEC-1999;
            BAYER AG (DE)
FEATURES    source
            1..3700
            /organism="Heliothis virescens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:7102"
            335..1825
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAC07500.1"
            /db_xref="GI:9996845"
            /translation="MGGRRARRSHLAAPAGLLLLCLLWPRGARGCYHEKRLHLLHLDH
            YNLERPVNESDPLQSLFGLTLMQIIDVDKKNQILITNLWLEWMDMLRWNTSDP
            GGKDLRVPNRLWPKDVLVYNSADGFDSTYPTNVVNRNGSLVYVPGIFKSTCKI
            DITWFPDDORCEMKGSWYDGYQLDLQLQDGGDISSFVFTNGEWELGVDPKNE
            IYNCPEPIDITFAVVIRKTLTYFFNLIVPCVLIASWALGFTLPPDSGKLSIG
            VTLLSLTFELNVAETMPATSDAVPLGLTYFNCIMPMVASSVSTLILINLHHRHAD
            THEMDSWIRCVFTLFWLVRMSRPGSATTPPARVPPDLELRERSSKSLANLVD
            IDDDFRHQAOQCCRYRGGENGAGLAHSCFVVDYELSLILKEIRVITDQMRKD
            DEDADISRDKWFAAMVVDRLCIIFTLTIIATLAVLSAPHIMVS"
            /db_xref="RENTREMBL:CAC07500"
            /translat="MGGRRARRSHLAAPAGLLLLCLLWPRGARGCYHEKRLHLLHLDH
            YNLERPVNESDPLQSLFGLTLMQIIDVDKKNQILITNLWLEWMDMLRWNTSDP
            GGKDLRVPNRLWPKDVLVYNSADGFDSTYPTNVVNRNGSLVYVPGIFKSTCKI
            DITWFPDDORCEMKGSWYDGYQLDLQLQDGGDISSFVFTNGEWELGVDPKNE
            IYNCPEPIDITFAVVIRKTLTYFFNLIVPCVLIASWALGFTLPPDSGKLSIG
            VTLLSLTFELNVAETMPATSDAVPLGLTYFNCIMPMVASSVSTLILINLHHRHAD
            THEMDSWIRCVFTLFWLVRMSRPGSATTPPARVPPDLELRERSSKSLANLVD
            IDDDFRHQAOQCCRYRGGENGAGLAHSCFVVDYELSLILKEIRVITDQMRKD
            DEDADISRDKWFAAMVVDRLCIIFTLTIIATLAVLSAPHIMVS"

ORIGIN
Alignment Scores:
Pred. No.:      5,99e-247      Length:      3700
Score:          2665.00      Matches:      496
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:      0
DB:              6              Gaps:        0

US-09-303-232-4 (1-496) x AX009612 (1-3700)

Qy      1  MetGlyArgAlaArgArgSerHisLeuAlaProAlaGlyLeuLeuLeuLeu 20
Db      335  ATGGGCGGCGGGCGCGCGCTCGACATTGGCGCGCGCGCGCGCTGCTGCTGCTG 394
Qy      21  CysLeuLeuTTPProArgGlyAlaArgCysGlyTyrHisGluLysArgLeuLeuHis 40
Db      395  TGCCTGCTCTGGCCGAGGGGGGCGACGCTGCGGGTACCAACGAGAGCGGCTACTGCACC 454
Qy      41  LeuLeuAspHisTyrAsnValLeuGluArgProValValAsnGluSerAspProLeuGln 60
Db      455  CTATTGGACCACTACAACGTACTGAGAGGCGCGCTCGTCAACGAGAGCGACCCGCTGCAG 514
Qy      61  LeuSerPheGlyLeuThrLeuMetGlnIleleAspValAspGluLysAsnGlnLeuLeu 80
```

```
Db      515  CTCCTCTTCGGCTCAGCTCATGCGATCATCGAGTGGACGAGAACACGACTTTTA 574
Qy      81  IleThrAsnIleTrpLeuLysLeuGluTrpAsnAspMetAsnLeuArgTrpAsnThrSer 100
Db      575  ATAACAAACATCTGGCTAAACTAGAGTGGATGATATGAACCTTGAGGTGGAACACTTCA 634
Qy      101  AspPheGlyGlyValLysAspLeuArgValProProHisArgLeuTrpLysProAspVal 120
Db      635  GATTTTCGGCGGGGTCAAAGATTTAAGAGTGCCACCCACAGACTATGGAACACGAGCGTC 694
Qy      121  LeuMetTyrAsnSerAlaAspGluGlyPheAspSerThrTyrProThrAsnValValVal 140
Db      695  CTTATGTCAACACGCGCGGACGAGGGTTCGACAGCAGCTATCCACGACAGCTGGTG 754
Qy      141  ArgAsnAsnGlySerCysLeuTyrValProProGlyIlePheLysSerThrCysLysIle 160
Db      755  CGGAACAACGGCTCGTGTCTGTACGTGCGCGCGCATCTTCAAGAGCACCTGCAAGATC 814
Qy      161  AspIleThrTrpPheProPheAspAspGlnArgCysGluMetLysPheGlySerTrpThr 180
Db      815  GACATCACCTGGTTCCTTCGACGACCAACGATGCGAGATGAAGTITGGCAGCTGACT 874
Qy      181  TyrAspGlyTyrGlnLeuAspLeuGlnLeuAspGluGlyGlyGlyAspIleSerSer 200
Db      875  TATGATGGTTATCAGTTGGATCTACAACTACAGATGAGGGGCGGAGATATAAGCAGT 934
Qy      201  PheValThrAsnGlyGluTrpGluLeuIleGlyValProGlyLysArgAsnGluIleTyr 220
Db      935  TTTGTACGAATGGCAATGGAGATTAAATAGAGTCCCGGCAAGCGCAACGAGATCTAC 994
Qy      221  TyrAsnCysCysProGluProTyrIleAspIleThrPheAlaValValIleArgAspLys 240
Db      995  TACAACCTGTTGTCGGAGGCATACATCGACATCACGTTTCGGGTGGTGAATCCGAGGAAA 1054
Qy      241  ThrLeuTyrTyrPhePheAsnLeuIleValProCysValLeuIleAlaSerMetAlaLeu 260
Db      1055  ACCTCTACTACTCTTCATCTGATCGTCCCTGGCTCATCGCTCCATCGCTCTA 1114
Qy      261  LeuGlyPheThrLeuProProAspSerGlyGluLysLeuSerLeuGlyValThrIleLeu 280
Db      1115  TTGGGGTTACCTTGGCTCCAGACTCCGAGAAAAGTTGTCTTATAGGTGTGACGATATTA 1174
Qy      281  LeuSerLeuThrValPheLeuAsnMetValAlaGluThrMetProAlaThrSerAspAla 300
Db      1175  CTGTCGTGTAGCGGTGTTCTCAACATGTCGGGAGACGATCCGAGCGAGCTCGACGCC 1234
Qy      301  ValProLeuLeuGlyThrTyrPheAsnCysIleMetPheMetValAlaSerSerValVal 320
Db      1235  GTGCCCTTGTTCGGCACCTACTTCACTGCATCATGTTTCATGTTGGTGGCTTCTCTCGTCGTC 1294
Qy      321  SerThrIleLeuLeuLeuAsnTyrHisHisArgHisAlaAspThrHisGluMetSerAsp 340
Db      1295  TCCACCATACTGATTCCTCAACTACCAACCGGACGACGACACTCAGAAATAGATGAT 1354
Qy      341  TrpIleArgCysValPheLeuTyrTrpLeuProTrpValLeuArgMetSerArgProGly 360
Db      1355  TGGATTCTGTCGTGTTCTCTTATTTGGCTCCGCTGGGTGCTGCCCATGTACGCCCCCGGC 1414
Qy      361  SerAlaThrThrProProProAlaArgValProProProProAspLeuGluLeuArgGlu 380
Db      1415  TCGGCGAGACGCGCGCGCGCGGTACCTCCGCGCGGACCTGGAGCTGGAGTCGCGAG 1474
Qy      381  ArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAspAspAspPheArgHisPro 400
Db      1475  CGCTCTCTCCAAGTCGCTCTAGCGAAGCGTCTGCATCGATCGATCGACGCTTCCGCGCACCCG 1534
Qy      401  GlnAlaGlnGlnProGlnCysArgTyrTyrArgGlyGlyGluGluAsnGlyValagly 420
Db      1535  CAACGCGAGACGCGCGCAATCTGCCGATCTACAGGGGGGGTGGAGAGATGGCGCGGGG 1594
Qy      421  LeuAlaAlaHisSerCysPheGlyValAspTyrGluLeuSerLeuLeuLeuLysGluIle 440
```


Db 1595 TTGGCGGCACAGTTGCTTGGTGTGCACTACGAGCTCTCCCTCATTTCTGAAGAGATT 1654
 QY 441 ArgValIleThrAspGluMetArgLysAspAspGluAspAlaAspIleSerArgAspTrp 460
 Db 1655 AGAGTCATCAGATCAGATCGCGAAGGACGACGAGGATCGGACATTTCCGCGCAGCTGG 1714
 QY 461 LysPheAlaAlaMetValValAspArgLysLeuLeuIlePheThrLeuPheThrIle 480
 Db 1715 AAGTTCCGCGCATCGTGTGGACAGACTGTGGCTTATATCTTTACCTGTTCAATC 1774
 QY 481 IleAlaThrLeuAlaValLeuLeuSerAlaProHisIleMetValSer 496
 Db 1775 ATCGCCACGCTAGCGTGTGTGTCGCGGCACACATCATGTGTGTCG 1822

RESULT 3

E58347
 LOCUS E58347 3701 bp DNA linear PAT 18-JUN-2001
 DEFINITION Nucleic acid encoding insect acetyl choline receptor subunit.

ACCESSION E58347

VERSION E58347.1 GI:13019346

KEYWORDS JP 2000023680-A/2.

SOURCE Heliothis virescens (tobacco budworm)

ORGANISM Heliothis virescens

REFERENCE Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea; Noctuidae; Heliothinae; Heliothis.

AUTHORS Martin, A., Nadja, E. and Thomas, S.

TITLE Nucleic acid encoding insect acetyl choline receptor subunit

JOURNAL Patent: JP 2000023680-A 2 25-JAN-2000;

COMMENT BAYER AG

OS Heliothis virescens

PN JP 2000023680-A/2

PD 25-JAN-2000

PF 26-APR-1999 JP 1999118159

PR 04-MAY-1998 DE 19819829.9

PI MARTIN ADAMUTSUUSUKI, NADJA ERASU, THOMAS SCHULTE PC

C12N15/09, A01K67/033, C07K14/705, C07K16/28, C12N1/21, C12N5/10, PC

C12Q1/68,

PC G01N33/15, G01N33/50// (C12N1/21, C12R1/19), C12N15/00, C12N5/00 CC

FH Key Location/Qualifiers

FT CDS (335). (1822).

1. .3701

/organism="Heliothis virescens"

/mol_type="genomic DNA"

/db_xref="taxon:7102"

FEATURES

source

ORIGIN

Alignment Scores:
 Pred. No.: 5.99e-247 Length: 3701
 Score: 2685.00 Matches: 496
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-09-303-232-4 (1-496) x E58347 (1-3701)

QY 1 MetGlyClyArgAlaArgArgSerHisLeuAlaAlaProAlaGlyLeuLeuLeuLeu 20
 Db 335 ATGGGCGGGCGGGCGCGCGCTGCACTTGGCGGCGCGCGCGCTGCTGCTGCTG 394
 QY 21 CysLeuLeuTrpProArgGlyAlaArgCysGlyTyrrHisGluLysArgLeuLeuHisHis 40
 Db 395 TGCTGCTCTGGCCGAGGGGGGACGCTGGGTACCCAGAGAGCGGCTACTGACAC 454
 QY 41 LeuLeuAspHisTyrrAsnValLeuGluArgProValValAsnGluSerAspProLeuGln 60
 Db 455 CTATTGACCACTACAACGCTACTGGAGAGCGCGCTGCTCAACGAGAGCGACCGCTG 514
 QY 61 LeuSerPheGlyLeuThrLeuMetGlnIleIleAspValAspGluLysAsnGlnLeuLeu 80

515 CTCCTCTCGGCTCACCTCATGAGATCATCGAGTGGAGAGAACAGCTTTTA 574
 QY 81 IleThrAsnIleTrpLeuLysLeuGluTrpAsnAspMetAsnLeuArgTrpAsnThrSer 100
 Db 575 ATAACAACATCTGGCTAAACTAGAGTGAATGATATGAACCTTGAGGTGGAACACTTCA 634
 QY 101 AspPheGlyGlyValLysAspLeuArgValProProHisArgLeuTrpLysProAspVal 120
 Db 635 GATTTCGCGGGGGTCAAGATTTAAGAGTTCACCCACAGACTATGGAACACGACGCTC 694
 QY 121 LeuMetTyrrAsnSerAlaAspGluGlyPheAspSerThrTyrrProThrAsnValVal 140
 Db 695 CTTATGTACAACAGCGCGGAGCGAAGGTTTCAGACGACGATTCACACGAGCTGGTGT 754
 QY 141 ArgAsnAsnGlySerCysLeuTyrrValProProGlyIlePheLysSerThrCysLysIle 160
 Db 755 CGGAACAACGGCTCGTGTCTGTACGTGCGCGCGGCATCTTCAAGAGCACCTGCAAGATC 814
 QY 161 AspIleThrTrpPheProPheAspAspGlnArgCysGluMetLysPheGlySerTrpThr 180
 Db 815 GACATCACCTGGTTCCTCCCTTCGACGACCAACGATCGGAGATGAAGTTTGGCAGCTG 874
 QY 181 TyrAspGlyTyrrGlnLeuAspLeuGlnLeuAspGluGlyGlyGlyAspIleSerSer 200
 Db 875 TATCATGGTTATCATGTTGGATCTACACTACAGATGAGGGGGCGGAGATATAGCACT 934
 QY 201 PheValThrAsnGlyGluTrpGluLeuIleGlyValProGlyLysArgAsnGluIleTyrr 220
 Db 935 TTTGTACGATGCGGAATGGGAGTTAATAGAGTCCCGCAAGCGCACGAGATCTAC 994
 QY 221 TyrAsnCysCysProGluProTyrrIleAspIleThrPheAlaValIleArgArgLys 240
 Db 995 TACAACCTGTTGTCGGAGGCCATACATGACATCATCGTTTCGGTGGTGTATCCGAGGAAA 1054
 QY 241 ThrLeuTyrrTyrrPhePheAsnLeuIleValProCysValLeuIleAlaSerMetAlaLeu 260
 Db 1055 ACGCTCTACTACTTCTTCAATCTGATCGTGGCCCTGGTGTCTCATCGCTCCATGGCTCTA 1114
 QY 261 LeuGlyPheThrLeuProProAspSerGlyGluLysLeuSerLeuGlyValThrIleLeu 280
 Db 1115 TTGGGGTTTCACTTCCCTCCAGACTCCGAGAGAAAAGTTGTCTTTAGGTGTGACATATTA 1174
 QY 281 LeuSerLeuThrValPheLeuAsnMetValAlaGluThrMetProAlaThrSerAspAla 300
 Db 1175 CTGTGTTGACGGTTCCTCAACATGTCGGAGACGATGCCAGGACGCTCGAGCGCC 1234
 QY 301 ValProLeuLeuGlyThrTyrrPheAsnCysIleMetPheMetValAlaSerSerValVal 320
 Db 1235 GTGCCCTTGTCTGGCACCTTCACTGTCATCATGTTTCATGGTGGCTTCTCTCGTGTCT 1294
 QY 321 SerThrIleLeuLeuLeuAsnTyrrHisArgHisAlaAspThrHisGluMetSerAsp 340
 Db 1295 TCCACCATATGATCTCAACTACACACCGGACGACACACTACGAAATGAGTGAT 1354
 QY 341 TrpIleArgCysValPheLeuTyrrTrpLeuProTrpValLeuArgMetSerArgProGly 360
 Db 1355 TGGATTCTTGGTGTCTTCTTTTGTGCTGCGGTGCTGCGCATGCAGGCCCGCCG 1414
 QY 361 SerAlaThrThrProProAlaArgValProProProAspLeuGluLeuArgGlu 380
 Db 1415 TCGGCGACGACGCGCGCGCGCGGTACTCTCCGCGCGCGACCTGGAGCTGCGCGAG 1474
 QY 381 ArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAspAspAspPheArgHisPro 400
 Db 1475 CGTCTCTCCAAAGTCGTCTCTAGCGAACGTGCTCGACATCGATGACGACTTCCGCGCAC 1534
 QY 401 GlnAlaGlnGlnProGlnCysCysArgTyrrTyrrArgGlyGlyGluGluAsnGlyValArg 420
 Db 1535 CAAGCGCAGCAGCGCGCAATGCTCCCGATATACTACAGGGGGGTGAGGAGATGGCGCGGG 1594
 QY 421 LeuAlaAlaHisSerCysPheGlyValAspTyrrGluLeuSerLeuLeuLeuLeuLeu 440

```

Db      1595 TTGGCGCGCACAGTTCGTCGGTGCAGTACGAGCTCCCTCATCTTCAAGGAGATT 1654
Qy      441 ArgVallleThrAspGlnMetArgLysAspAspGluAspAlaAspIleSerArgAspTyr 460
Db      1655 AGAGTCATCATCAGATCGCGCAGGACGACGAGATCGGACATTCGCGCGACTGG 1714
Qy      461 LysPheAlaAlaValValAspArgLeuCysLeuIleIlePheThrLeuPheThrIle 480
Db      1715 AAGTTCCCGCCGACGCGGCGGACAGACTGCGCTTATTATCTTTACCTGTTTCAACATC 1774
Qy      481 IleAlaThrLeuAlaValLeuLeuSerAlaProHisIleMetValSer 496
Db      1775 ATGCCACGCTAGCGGCTGCTCGCGGCCACACATCATGCTGTCG 1822

RESULT 4
AF272778      2907 bp      mRNA      linear      INV 29-APR-2002
LOCUS      Drosophila melanogaster nicotinic acetylcholine receptor Dalpha5
DEFINITION      subunit (nAcRalpha-34E) mRNA, nAcRalpha-34E-A allele, complete cds.
ACCESSION      AF272778
VERSION      AF272778.1 GI:20152839
KEYWORDS
SOURCE      Drosophila melanogaster (fruit fly)
ORGANISM      Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE      1 (bases 1 to 2907)
AUTHORS      Grauso, M., Reenan, R.A., Culetto, E. and Sattelle, D.B.
TITLE      Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
Dalpha5, Dalpha6 and Dalpha7, in Drosophila melanogaster Identify a
New and Highly Conserved Target of Adenosine Deaminase Acting on
RNA-Mediated A-to-I Pre-mRNA Editing
Genetics 160 (4), 1519-1533 (2002)
JOURNAL
MEDLINE      21969411
PUBMED      11973307
REFERENCE      2 (bases 1 to 2907)
AUTHORS      Grauso, M. and Sattelle, D.B.
TITLE      Direct Submission
JOURNAL
FEATURES
source      1. .2907
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/chromosome="2"
/map="2L; 34E4-5"
/dev_stage="embryo"
2. .2907
/gene="nAcRalpha-34E"
/allele="A"
289. .2712
/gene="nAcRalpha-34E"
/notes="ion channel; neurotransmitter transmembrane
receptor"
/codon_start=1
/product="nicotinic acetylcholine receptor Dalpha5
subunit"
/protein_id="AAMI3390.1"
/db_xref="GI:20152840"
/translators="MKNAOLKLETVDDDELAVRLACSSNFSSSSSTSTSSNOBH
NQUTLQPSLSTKHHNSINASEQNSQQEPASKDEEDVANHGNSDQQLQLODSS
NMUSPTKAAATAGDEATQQPTNIELCKARKQRRLRRRRKPATNETDRIKQOOL
SMFPFKTSTYSTAATSCPTATMQCRASDNFESIPISRHDRVSTATFAWVLIH
VLOVLLVLOQWOLHVQORSVLLFRRIAASATFIAPISVLGFAAOLKNSSSSSSSNS
NNSSTOGLNKHWSWTFLLYLNLAKVCLAGYHEKRLHLLDLPDNTLIERPVINES
DPLQLSGLTLMQIIDVDEKNQLLVTVMLKLEWDMNLEWNTSDYGVKDLRIPHR
IKRFDVLMNSADRGFGTGTQTNVVVRNNGSLVPPGIFKSTKIDITWFPDDQRC
EMKFGSYDGFQDLQLOQDETGDSSVINGEWELLGPGKRNELIYNCCPEPYID
ITFAIIRRLTYFFNLIIPCVLIASMALIGFTLPDPSGKLSGLVITLLSLTVLIN
MVAETPMATSDAVPLGTYFNCIMFVASSVSVTLILNYHHRNADTHEMSEWRIVF
LCWLPWILRMSRPGRLILEFPTTPCSDTSSERKHQILSDVELKERSKSLIANVLID

```

variation

variation

variation

variation

variation

variation

variation

ORIGIN

Alignment Scores:

```

Pred. No.:      4,47e-189      Length:      2907
Score:          2064.50      Matches:      396
Percent Similarity:      82.95%      Conservative:      32
Best Local Similarity:      76.74%      Mismatches:      49
Query Match:      77.47%      Indels:      39
DB:              3      Gaps:      6

```

US-09-303-232-4 (1-496) x AF272778 (1-2907)

Qy 17 LeuLeuLeuLeuCysLeuLeuTyrProArgGlyAlaArgCysGlyTyrHisGluLysArg 36

Db 1171 TTATTGATATATTGAATTTATCTGCTAAAGTTTTCGCTAGCAGGATATCATGAAAAGAGA 1230

Qy 37 LeuLeuHisHisLeuLeuAspHisTyrAsnValLeuGluArgProValValAsnGluSer 56

Db 1231 CTGTTACACCATCTTTTGGATCCTTATAATACACTAGAACGTCCTCGTCTCAATGAATCG 1290

Qy 57 AspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleLeuAspValAspGluLys 76

Db 1291 GACCCGTTACAATTAGCTTTGGTTTAACTTAACTTAATCAAAATTCGATGTGGACGAGAAA 1350

Qy 77 AsnGlnLeuLeuIleThrAsnIleTyrLeuLysLeuGluTyrAsnAspMetAsnLeuArg 96

Db 1351 AATCAATTTGATGTCATTAATGTGGTTAAAACTGAGTGGAAACGACATGAATCTCCGC 1410

Qy 97 TrpAsnThrSerAspPheGlyGlyValLysAspLeuArgValProProHisArgLeuTyr 116

Db 1411 TGGAAACACTCCGACTATGCGGAGTTAAGATCTCGAATACCGCGCATCGCATCTGG 1470

Qy 117 LysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspSerThrTyrProThr 136

Db 1471 AAGCCGACGCTGCTGATGTACAACAGTCGCGATGAGGGATTTGACGGCACCTACCACAGC 1530

137 AsnValValValArgAsnAsnGlySerCysLeuTyrValProProGlyIlePheLysSer 156
1531 AACGTGGTGGTGGCGAACAACGGCTCGTGTCTATACGTTCCGCGGGGATCTTCAAGTCG 1590
157 ThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCysGluMetLysPhe 176
1591 ACGTGCAGATCGACATCACGTGGTTCCTTCGATGACACGCGGTGGAGATGAAGTTC 1650
177 GlySerTrpThrTyrAspGlyTyrGlnLeuAspLeuGlnAspGluGlyGly 196
1651 GGCAGTTGAGACTGACGCGAATCCAGCTGGATTTACAAATACAGATGAACACGCGCGGT 1710
197 AspIleSerSerPheValThrAsnGlyGluTrpGluLeuIleGlyValProGlyLysArg 216
1711 GATATCAGCAGTACGTGCTCAACGCGGAGTGGGAACCTACTGGGTGTCGCCGCAACAGT 1770
217 AsnGluIleTyrThrAsnCysCysProGluProTyrIleAspIleThrPheAlaValVal 236
1771 AACGAGATCTATTACAACCTGCTGCCGGAACCTATATAGACATCACCTTCGCCATCATC 1830
237 IleArgArgLysThrLeuTyrTyrPhePheAsnLeuIleValProCysValLeuIleAla 256
1831 ATCCGCCGACGAAACACTGTACTATTCTTCAACCTGATCATACCTTGTGTACTGATGCC 1890
257 SerMetAlaLeuLeuGlyPheThrLeuProAspSerGlyGluLysLeuSerLeuGly 276
1891 TCCATGGCTTGTCTCGGATTCACCTCGCCGACGATTCGGGTGAAAAATATTCGCTGGT 1950
277 ValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAlaGluThrMetProAla 296
1951 GTTACCATCTTGTCTCTGCTCAGCGTGTCTGTAATATGTTGGTGGAGACAATCGCGCT 2010
297 ThrSerAspAlaValProLeuLeuGlyThrTyrPheAsnCysIleMetPheMetValAla 316
2011 ACTTCCGATGCGGTGCCATTCGCTGGGTACATATTTCAATTCATATGTTTATGGTAGCT 2070
317 SerSerValValSerThrIleLeuLeuAsnTyrHisArgHisAlaAspThrHis 336
2071 TCATCCGTTGTGTCAACGATTTAAATATTAATATATCATCATCGAATGCTGATACGCAC 2130
337 GluMetSerAspTrpIleArgCysValPheLeuTyrTrpLeuProTrpValLeuArgMet 356
2131 GAAATGTCGAATGGAACGATCGTGTGTTTGTGCTGGCTGCCATGATATTCGAATG 2190
357 SerArgProGlySerAlaThrProProProAlaArgValProProProProProProPro 374
2191 AGTCCGCCGAGA-----CGACCGCTGATCTCTAGAGTCCGACACACCGCTGTTCG 2241
375 -----AspLeuGluLeuArgGluArgSer 382
2242 GACACATCTCCGAGCGGAACGACGACGATCTCTCCGACGTTGAGCTGAAAGACGCTCG 2301
383 SerLysSerLeuLeuAlaAsnValLeuAspIleAspAspAspPheArgHis----- 399
2302 TCGAAATCGCTGCTGGCCAGCTACTAGACATCATGATGATCTCCGCGACAAATGTGCG 2361
400 -----ProGlnAlaGlnProGlnCysCysArgTyrTyrArg----- 412
2362 CCCATGACCGCCGCGGAACACTGCCACACACCGCGCTTCTATCGACGCTTATGGA 2421
413 GlyGlyGluGluAsnGlyAlaGly-----LeuAla 422
2422 CAAGCGCAGCTGACGATTCGCGCCCAATTTGGCAGCACCCGAAATGCGGATGCGCTCAC 2481
423 AlaHisSerCysPhe-----GlyValAspTyrGluLeuSerLeuIleLeuLysGluIle 440
2482 CATCATGCTGCATCAATCATCATCAATGAATAGATTTAACTTTAAAGGAAT 2541
441 ArgValIleThrAspGlnMetArgLysAspAspGluAspAlaAspIleSerArgAspTrp 460
2542 CGCTTTATACTGATCAGCTAGCTAAAGATGACGAGTGAATGACATTCGCAATGATGG 2601
461 LysPheAlaAlaMetValValAspArgLeuLeuLeuIlePheThrLeuPheThrIle 480

1662 AAAATTGACGATGCTGCTGACAGACTGTCCTATCATATTCACATGTTCCCAATA 2661
481 IleAlaThrLeuAlaValLeuLeuSerAlaProHisIleMetValSer 496
2662 TTAGCCACAATAGCTGTAATACTATCATCAGCACCATATTAATGTTCTCG 2709
RESULT 5
DME554210 1683 bp mRNA linear INV 01-APR-2003
LOCUS Drosophila melanogaster mRNA for nicotinic acetylcholine receptor
DEFINITION subunit Dalpaph7 (nAcRalpha-18C gene).
ACCESSION AJ554210
VERSION AJ554210.1 GI:29466436
KEYWORDS nAcRalpha-18C gene; nicotinic acetylcholine receptor subunit
Dalpaph7.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1683)
AUTHORS Millar,N.S.
TITLE Direct Submission
JOURNAL Submitted (31-MAR-2003) Millar N.S., Department of Pharmacology,
University College London, Gower Street, London, WC1E 6BT, UNITED
KINGDOM
FEATURES
source Location/Qualifiers
1..1683 /organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
1..1683 /gene="nAcRalpha-18C"
26..1654 /gene="nAcRalpha-18C"
/codon_start=1
/product="nicotinic acetylcholine receptor subunit
Dalpaph7"
/protein_id="CAD86936.1"
/db_xref="GI:29466437"
/translation="MSFPQPHSLPEATANGRMVLYGGLLIMIPACAAAPHEKRLRH
ALLDYNLSRPVYVNSDPLQLSGLTLMQIIDVENKQLLITNWLKWNMDNLRW
NSSEFGVGRDLRIPPHRLWKPDLVYNSADEGDTATVTVNVRNGSLVYVPGLIFK
STCKIDITWPFDDORCEMGKGSWTYGFQDLQLODEAGDIISSFTINGSDLLGVP
GKNEIYYNCCPXYIDITFALLRRKTYFFNLYVFCVLIASKALLGFTLPDPSGE
KLSGLTLLSLVFLNMVAETMPATSDAVLLGKIFNCFMFWASSVSTLILVNIYH
HRNPDTHEMSEWIRVIFYWLPCLLRMRQGVGECPPPPSSSSSSASGKKQOIQN
VELKERSKLLANVLDDDDPCNHRCSATLPHQTYTYRQYRQDGGSGVGPQA
GPVVDGRLEHAISHTCLTSSAEYELALILKELRWITELQKKEDETSITRDWKEAAMV
VDRLCLITFTLTIIATLAVLPSAPHFIVSGVRG"
26..131
/gene="nAcRalpha-18C"
132..1651
/gene="nAcRalpha-18C"
/product="nicotinic acetylcholine receptor subunit
Dalpaph7"
sig_peptide
mat_peptide
ORIGIN
Alignment Scores:
Pred. NO.: 1.01e-187 Length: 1683
Score: 2047.00 Matches: 399
Percent Similarity: 81.16% Conservative: 36
Best Local Similarity: 74.44% Mismatches: 47
Query Match: 76.81% Indels: 54
DB: 3 Gaps: 9
US-09-303-232-4 (1-496) x DME554210 (1-1683)
Qy 2 GlyGlyArgAlaArgArgSerHisLeuAlaProAlaGlyLeuLeuLeuLeuCys 21
Db 71 GGTGGCAGATG-----CTGCTATGCGCTGGAGCTTTTAATTATGATA--- 115

QY 22 LeuLeuTrpProArgGlyAlaArgCysGlyTyrHisGluLysArgLeuLeuHisHisLeu 41
Db 116 -----CCGGCTTGTGGCTGGACCCCATGAGAGCGGCTACTCCAGCCCTT 163
QY 42 LeuAspHisTyrAsnValLeuGluArgProValValAsnGlnSerAspProLeuGlnLeu 61
Db 164 CTGGACAACCTACACAGCGCTCGAGCGTCCGGTGGTCAATGAATCCGATCCATTGCAACTG 223
QY 62 SerPheGlyLeuThrLeuMetGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 81
Db 224 AGCTTCGGACTTAAACACTCATGTCAGATTATCATGTCGACGAAAGAAATCAACTGCTTATA 283
QY 82 ThrAsnLeuTrpLeuLysLeuGluTrpAsnAspMetAsnLeuArgTyrAsnThrSerAsp 101
Db 284 ACGATATTTGGCTCAAAATGGAATGGAACGATATGAATCTTCGATGGAATTCAGATGGAG 343
QY 102 PheGlyGlyValLysAspLeuArgValProProHisArgLeuTrpLysProAspValLeu 121
Db 344 TTCGGTGGTGTGGCGGATCTGCGAATTCGCGCACATCGCTATGGAACCGGATGTACTG 403
QY 122 MetTyrAsnSerAlaAspGluGlyPheAspSerThrTyrProThrAsnValValArg 141
Db 404 ATGTACAAACAGTCGCGACGAGGGCTTCGATGGAACTGACGACCAAAATGTGGTTCGCTCGC 463
QY 142 AsnAsnGlySerCysLeuTyrValProProGlyTyrPheLysSerThrCysLysIleAsp 161
Db 464 AATATCGGGAGCTGTCTGTACGTACCGCCAGGTATATTTAAGTCAACGTGTAAATCGAC 523
QY 162 IleThrTrpPheProPheAspAspGlnArgCysGluMetLysPheGlySerThrTyr 181
Db 524 ATTACGTGGTTCCTTCGACGATCAGAGATGTGAAATGAAATTTGGTTCGTGGACCTAC 583
QY 182 AspGlyTyrGlnLeuAspLeuGlnLeuGlnAspGluGlyGlyAspIleSerSerPhe 201
Db 584 GATGGGTTTCAGTTGGACCTTCAGTTGCAGACAAAGCTGTGGCGACATTTCTAGCTTT 643
QY 202 ValThrAsnGlyGluTrpGluLeuIleGlyValProGlyLysArgAsnGluLeuTyrTyr 221
Db 644 ATAACCAATGGCGATGGGACTTGTAGGTGTGCCGTAAACGAAATGAAATCTACTAT 703
QY 222 AsnCysCysProGluProTyrIleAspIleThrPheAlaValValIleArgAtgLyThr 241
Db 704 AATTGCTGCCAGAACCTTATATTGACATAACATTCGCCATTTGTATGAAAGCGCAAAACG 763
QY 242 LeuTyrTyrPhePheAsnLeuLeuValProCysValLeuLeuIleAlaSerMetAlaLeuLeu 261
Db 764 TTGTACTATTTTTCATCTCTGATTTGTCGGTGTACTGATCGCTCCATGGCAGTCTGTA 823
QY 262 GlyPheThrLeuProProAspSerGlyLysLeuSerLeuGlyValThrIleLeuLeu 281
Db 824 GGGTTTACACTGCCACCAGATTCTGGTGAAAGCTTTTCGCTTGAGTTTACAATTTCTATA 883
QY 282 SerLeuThrValPheLeuAsnMetValAlaGluThrMetProAlaThrSerAspAlaVal 301
Db 884 TCGCTTACAGTCTTCCTCAACATGTTGGCGAAACAATGCGGACCTCCGATCGGTA 943
QY 302 ProLeuLeuGlyThrTyrPheAsnCysIleMetPheMetValAlaSerSerValSer 321
Db 944 CCGCTGTCGGTAGTATTTCAATTTGATTTATGTTTATGTTGGCTCATCATGTTGTGTA 1003
QY 322 ThrIleLeuIleLeuAsnTyrHisArgHisAlaAspThrHisGluMetSerAspTrp 341
Db 1004 ACCATCTTGTCTCAATTTATCATATGATATCCATAGATATCCATAGCAATGAATGAGTGAATGG 1063
QY 342 IleArgCysValPheLeuTyrTrpLeuProTrpValLeuArgMetSerArgProGlySer 361
Db 1064 ATAAGAGTAATATCTTTTATTTGTTTACCTTTCATATTTGCTATGTCATGCAAGACCGGACAG 1123
QY 362 AlaThrThrProProAlaArgValProProPro----- 374
Db 1124 GTT-----GGCTACGAATGTCCGCGCGCGCCCTTCTTCTCGAGTTCCTCGCA 1171
QY 375 -----AspLeuGluLeuArgGluArgSerSerLysSer 385

Db 1172 TCCGGCGAGAGAGAACAGATCCAAACCGTTCAGCTCAAGGAGAGTCTCTCAAGTCT 1231
QY 386 LeuLeuAlaAsnValLeuAspIleAspAspPheArg-----HisProGlnAla--- 402
Db 1232 CTGCTGGCCCAATGTGCTCGATATAGACGATGATTTCCGATGCAATCATCGATGTGCCAGC 1291
QY 403 -----GlnGlnProGlnCysCysArg---TyrTyrArgGlyGlyGluGluAsn 417
Db 1292 GCGACTTTTGCCCCACCACGACCATATATACAGGACGATGTACAGGCAAGGAGATGACGCG 1351
QY 418 GlyAlaGly-----LeuAla 422
Db 1352 AGCGTGGGACCGGTGGGACCGAGCTGTCAGTGTGGACGCGGCGCTTTGACAGAGGCCATT 1411
QY 423 AlaHisSerCysPhe-----GlyValAspTyrGluLeuSerLeuLeuLeuGluLeu 440
Db 1412 TCCACACCTGTCTGACATCTCTCGGAGTACGAACCTGGCGCTGATACTCAAGAGAGCTG 1471
QY 441 ArgValIleThrAspGlnMetArgLysAspAspGluAspAlaAspIleSerArgAspTrp 460
Db 1472 CGTTGGTAACAGAACAGCTCAAAAAAGAGACGAAACAAGCGACATTACGCGAGATTGG 1531
QY 461 LysPheAlaAlaMetValValAspArgLeuCysLeuIleIlePheThrLeuPheThrIle 480
Db 1532 AATTTCTGCCATGGTGTGCGATCGTGTGTCCTTATTTTCACCTTGTACTATT 1591
QY 481 IleAlaThrLeuAlaValLeuLeuSerAlaProHisIleMetValSer 496
Db 1592 ATAGCAACCTCGCTGTACTTCTTCAGCGCGCATTTTCATTGTGAGT 1639

RESULT 6

AY036613

LOCUS

DEFINITION

Drosophila melanogaster nicotinic acetylcholine receptor Dalphas subunit (nAcRalpha34E) mRNA, nAcRalpha34E-B allele, complete cds, alternatively spliced.

ACCESSION

AY036613

VERSION

AY036613.1

GI:20340268

SOURCE

Drosophila melanogaster (fruit fly)

ORGANISM

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

1 (bases 1 to 2834)

Gauso, M., Reenan, R.A., Culetto, E. and Sattelle, D.B.

Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes, Dalphas5, Dalphas6 and Dalphas7, in Drosophila melanogaster Identify a New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-Mediated A-to-I Pre-mRNA Editing

Genetics 160 (4), 1519-1533 (2002)

JOURNAL

MEDLINE

21969411

PUBMED

11973307

REFERENCE

2 (bases 1 to 2834)

Gauso, M. and Sattelle, D.B.

Direct Submission

Submitted (26-MAY-2001) MRC-FGU Human Anatomy and Genetics, University of Oxford, South Parks Road, Oxford OX1 3QX, UK

JOURNAL

FEATURES

source

1. .2834

/organism="Drosophila melanogaster"

/mol_type="mRNA"

/db_xref="taxon:7227"

/chromosome="2"

/map="34E4-34E5"

/dev_stage="embryo"

1. .2834

/gene="nAcRalpha34E"

/allele="B"

289. .1737

/gene="nAcRalpha34E"

/note="ion channel; neurotransmitter transmembrane

QY	217	AsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThrPheAlaValVal	236
Db	1854	AACGAGATCTATTACAACCTGCTCCCGGAACCCTATATAGACATCATCTTCGCCATCATC	1913
QY	237	IleArgArgLysThrLeuTyrTyrPhePheAsnLeulleValProCysValLeulleAala	256
Db	1914	ATCCGGCGAGAACACTGTACTATTCTCCAACCTGTATCATCTTGTTACTGATTGCC	1973
QY	257	SerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeuSerLeuGly	276
Db	1974	TCCATGGCCTTGCTCGGATCACCTGCCGCAGATTCGGGTGAATAATATATCGCTGGGT	2033
QY	277	ValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAlaGluThrMetProAla	296
Db	2034	GTTACCATCTTGCTCTCGCTAGCCGTGTTTCTGAATATGCTTCCGAGACAAATGCCGCT	2093
QY	297	ThrSerAspAlaValProLeuLeuGlyThrTyrPheAsnCysIleMetPheMetValAla	316
Db	2094	ACTTCCGATCGCGTGCCATTG-----	2114
QY	317	SerSerValValSerThrIleLeulleLeuAsnTyrHisIсарgHisAlaAspThrHis	336
Db	2114	-----	2114
QY	337	GluMetSerAspTrpIleArgCysValPheLeuTyrTrpLeuProTrpValLeuArgMet	356
Db	2115	-----TGGATACGATCGTGTGTTTTGCTGCTGCCATGATTAATTCGCAATG	2162
QY	357	SerArgProGlySerAlaThrThrProProAlaArgValProProProPro-----	374
Db	2163	AGTCGCCACGGA-----CGACGCTGATCCTTAGAGTTCCCGACCAACGCCCTGTTCG	2213
QY	375	-----AspLeuGluLeuArgGluArgSer	392
Db	2214	GACACATCTCCGAGCGGAAGCACAGATACTCTCCGACGTTGAGCTGAAGAAGCGCTCG	2273
QY	383	SerLysSerLeuLeuAlaAsnValLeuAspIleAspAspPheArgHis-----	399
Db	2274	TCGAATCTGCTGCGCCAAAGTACTAGACATCGATGACATTCGGCACAAATGTGCG	2333
QY	400	-----ProGlnAlaGlnProGlnCysCysArgTyrTyrArg-----	412
Db	2334	CCCATGACCCCCGGCGGAACACTGCCACACACCCCGCTTCTATCGACGTTTATGGA	2393
QY	413	GlyGlyGluGluAsnGlyAlaGly-----LeuAla	422
Db	2394	CRAAGCGCATGGCGAGCATTTGGCCAAATGGCAGCACCCGAATGGCGGTGCGTCACC	2453
QY	423	AlaHisSerCysPhe-----GlyValAspTyrGluLeuSerLeulleLysGluIle	440
Db	2454	CATCATACGTGCATCAAATCATCACTGAATATGAATTAGTGTAACTTTAAAGGAAAT	2513
QY	441	ArgValIleThrAspGlnMetArgLysAspAspGluAspAlaAspIleSerArgAspTrp	460
Db	2514	CGCTTTATAACTGATCAGTACGTAAAGATGACGAGTGCAAATGACATTCGCAATGATGG	2573
QY	461	LysPheAlaAlaMetValValAspArgLeuCysLeuIleIlePheThrLeuPheThrIle	480
Db	2574	AAATTGCACTATGCTGTGACAGACTGTGCTTATCATATTCATATTCATATTCGCAATA	2633
QY	481	IleAlaThrLeuAlaValLeuLeuSerAlaProHisIleMetValSer	496
Db	2634	TTAGCCACAATAGCTGTACTACTATCGGCACACATATATTGTTCTCG	2681
RESULT 8			
AX009610			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Drosophila melanogaster (fruit fly)			
Sequence 1 from Patent EP0962528.			
AX009610			
AX009610.1 GI:9996842			
DNA linear PAT 06-SEP-2000			
2886 bp			


```
Db 2634 TTAGCCACATAGCTGTACTACTATCGGACCAACATATTATTGTCTCG 2681
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
RESULT 9
AF143847 3029 bp mRNA linear INV 27-MAY-1999
LOCUS Heliothis virescens putative nicotinic acetylcholine receptor alpha
DEFINITION 7-2 subunit mRNA, complete cds.
ACCESSION AF143847
VERSION AF143847.1 GI:4895006
KEYWORDS
SOURCE
ORGANISM Heliothis virescens (tobacco budworm)
Bukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuoidea; Noctuidae; Heliothinae; Heliothis.
REFERENCE 1 (bases 1 to 3029)
AUTHORS Schulte,T., Oellers,N. and Adamczewski,M.
TITLE Putative alpha subunits of insect nicotinic acetylcholine receptors
more similar to vertebrate alpha 7 subunits and C. elegans Ce21
than to other insect nicotinic acetylcholine receptor alpha
subunits
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3029)
AUTHORS Schulte,T., Oellers,N. and Adamczewski,M.
TITLE Direct Submission
JOURNAL Submitted (19-APR-1999) ZF-BTB, Bayer AG, Bldg. Q 18,, Leverkusen
51368, Germany
FEATURES
source
1..3029
/organism="Heliothis virescens"
/mol_type="mRNA"
/db_xref="taxon:7102"
95..1600
/notes="hvnachra7-2"
/codon_start=1
/product="putative nicotinic acetylcholine receptor alpha
7-2 subunit"
/protein_id="A032698.1"
/db_xref="GI:4895007"
/translations="MAPMLAALALLPVSEQGPHEKRLNALLANYNTPRVANE
SELEVRFGTLQIIIDVDEKNOLLITNIWLSLEWINDYLRWNSYGVVKDLRIIPN
KLMQDVLMYDNGEDFQYQTNVVRSGCLYVPBGIFKSTCKMDIAWPFDDQH
CDMKFSGWTYDNGQLDLVKADEAGDLSDFITNGEWYLIGMPGKNITFYACCPBYV
DVTFTIMRRRTIYFENLIVPCVLSSMALLGFTLPDPSGEKLTIGVILLSLTVL
NLVATLQPQSDAIPLLGTYFCNIMFWASSVLTIVLNLNHHRTADIHMPQWIKSV
FLQWLPLWLMSPGKKITRKTIMMTRMELEKRSKLLANLIDDDDFRHGFP
PNSASTGNLPGCSIFRTDFRSFVRPSTMEDVGGGLSHHREHLILRELQFITA
RMKKADEAEELISDMKFAAMVDRFCLFVFTLTIIATVAVLISAPHIIVQ"
ORIGIN
Alignment Scores:
Pred. No.: 2,31e-164 Length: 3029
Score: 1809.00 Matches: 352
Percent Similarity: 77.52% Conservative: 55
Best Local Similarity: 67.05% Mismatches: 78
Query Match: 67.88% Indels: 40
DB: 3 Gaps: 8
US-09-303-232-4 (1-496) x AF143847 (1-3029)
QY 5 AlaArgSerHisLeuAlaProAlaGlyLeuLeu----- 17
Db 38 GCGCGCGCGGTGACTACGCAGCGGTGCGGGGTGTGTATGCCGCGCGTATG 97
QY 18 -----LeuLeuLeuCysLeuLeuTrpProArgGlyAlaArgCysGly 31
Db 98 GCGCCATGTTGGCGGCTTGGCGCTGCTGGCTTGGTGGCGGTATCGGAGCAA---GGT 154
QY 32 TyrHisGluLysArgLeuLeuHisHisLeuLeuAspHisTyrAsnValLeuGluArgPro 51
Db 155 CCTCACGAGAGAGACTCCTCTGAACGCGTGTGTCGCGAACTACAAACACCTCGGACCG 214
52 ValValAsnGluSerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleIle 71
215 GTGGCCAAACGAGAGCGAACCCGTAGAGGTTCGGCTTCAGCTTCGACCAATCATT 274
QY 72 AspValAspGluLysAsnGlnLeuLeuIleThrAsnIleTrpLeuLysLeuGluTrpAsn 91
Db 275 GACGTGGACGAGAGAAGATCAACTACTATTATACCAATATATATGGCTGTGTTGGAGTGAAT 334
QY 92 AspMetAsnLeuArgTrpAsnThrSerAspPheGlyGlyValLysAspLeuArgValPro 111
Db 335 GACTACAACTGAGGTGGAACGACGAGTATATGGCGGGTCAAGGACCTCAGGATCAGC 394
QY 112 ProHisArgLeuTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAsp 131
Db 395 CCACAAAGTGTGGAAAGCGGACGCTTATATATAATAGTCTGACGAGGGTATTGAC 454
QY 132 SerThrTyrProThrAsnValValArgAsnAsnGlySerCysLeuTyrValProPro 151
Db 455 GGGACCTTACCAGACCAACGAGTGGTGTGACAGAGCGGGCGAGTTGCCGTGACGTGCCACCT 514
QY 152 GlyIlePheLysSerThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArg 171
Db 515 GGCATATTCGAAGACACATGCAAGATGGACATCGCGTGGTTCCTTCGACGACCAACAC 574
QY 172 CysGluMetLysPheGlySerTrpThrTyrAspGlyTyrGlnLeuAspLeuGln 191
Db 575 TGTGATATGAAGTTCGGTAGCTGCACATATGACGGCAATCAGTTGGATCTGGTCTGCTAAAA 634
QY 192 AspGluGlyGlyGlyAspIleSerSerPheValThrAsnGlyGluTrpGluLeuIleGly 211
Db 635 GATGAGCGAGCGCGCATCTATCGGACTTCAATAACAAATGGGAGTGTATCTAATAGGA 694
QY 212 ValProGlyLysArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIle 231
Db 695 ATGCCAGGCAAAAAGAACACATAACATACGCGTGTGTCGCCCGACGCTACGTCGACGCT 754
QY 232 ThrPheAlaValValIleArgArgLysThrLeuTyrTyrPhePheAsnLeuIleValPro 251
Db 755 ACCTTCACCATCATGATGAAGAACGAAACCTTGTACTTCTTCAACCTGATCGTCCCG 814
QY 252 CysValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGlu 271
Db 815 TGGGTGCTGATCTCATCGATGGGACTCTCGGCTTCACATGCCACAGACTCCGGAGAG 874
QY 272 LysLeuSerLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAla 291
Db 875 AAACCTCACACTTGGAGTCACTATCTTCTATCGCTGACGCTGTTCCTCAACCTGGTAGCC 934
QY 292 GluThrMetProAlaThrSerAspAlaValProLeuLeuGlyThrTyrPheAsnCysIle 311
Db 935 GAGACCTTGCACAGGTCTCCGACGCTATCCCTCTGTGGAGAGTACTTCAATGTCATC 994
QY 312 MetPheMetValAlaSerSerValValSerThrIleLeuIleLeuAsnTyrHisHisArg 331
Db 995 ATGTTTCATGTTAGCGTCTGTGGTGTACTGACTGTGGTGTGTACTCAATATACCACCATCGA 1054
QY 332 HisAlaAspThrHisGluMetSerAspTrpIleArgCysValPheLeuTyrTrpLeuPro 351
Db 1055 ACAGCTGATATACATGAAATGCCACAGTGGATAAAATCACTATTCCTACAAATGGTGGCCA 1114
QY 352 TrpValLeuArgMetSerArgProGlySerAlaThrThrProProAlaArgVal--- 370
Db 1115 TGGATCTCGGAATGTCGAGGCCAGGAGAGAGATCACAGGAAGACTATATGATGAAC 1174
QY 371 ProProProProLeuLeuLeuArgGluArgSerSerLysSerLeuLeuAlaAsnVal 390
Db 1175 ACCGAGGATGAGGAGCTGGAACTGAAGGAGAGGTCTGTCGAGTCTTGTGTCGGCGAATGTT 1234
QY 391 LeuAspIleAspAspPheArgHis-----ProGlnAlaGlnGln----- 404
Db 1235 CTAGATATTGATGATGACTTCAGACACGCGCCCTCCGCTCTTAACAGTACTGCTCGCTCGACC 1294
QY 405 -----ProGlnCysCysArgTyr-----TyrArgGlyGly----- 414
```

Db	1295	GGGAATTGGGACCTGGGTGCTCAATATCCGCACGGATTCCGTCGGTGGTTCGTC	CGT	1354
			:::	
Qy	415	-----GluGluAsnGlyAlaGlyLeuAlaAlaHisSerCysPheGlyValAsp		430
			:::	
Db	1355	CGGTCCACGATGGAAGACGTGGCGCGGCTGGGTAGCCAC	-----CAT	1399
Qy	431	TyrGluLeuSerLeuLeuLeuLysGluIleArgValIleThrAspGlnMetArgLysAsp		450
			:::	
Db	1400	CGCGAGCTGCACCTCATCTAGAGAGCTCAGTTTCATCAGCCAGGATGAAGAAGCT		1459
Qy	451	AspGluAspAlaAspIleSerArgAspTrpLysPheAlaAlaMetValValAspArgLeu		470
			:::	
Db	1460	GATGAGGAACCCAGAGCTGATCAGCGACTGGAAGTTTGCTCGCATGGTTGTGTAGT	TTT	1519
Qy	471	CysLeuIleIlePheThrLeuPheThrIleAlaThrLeuAlaValLeuLeuSerAla		490
			:::	
Db	1520	TGCCTGTTGGTGGTTCACATCTTTTCACAAUCATCGCGACAGTAGCTGCTCCTGTTATCGGCA		1579
Qy	491	ProHisIleMetVal		495
			:::	
Db	1580	CCGCATATCATCGTG		1594

RESULT 10	
E58348	
LOCUS	E58348
DEFINITION	Nucleic acid encoding insect acetyl choline receptor subunit.
ACCESSION	E58348
VERSION	E58348.1 GI:13019347
KEYWORDS	JP 2000023680-A/3.
SOURCE	Heliothis virescens (tobacco budworm)
ORGANISM	Heliothis virescens Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea; Noctuidae; Heliothinae; Heliothis. 1 (bases 1 to 3109) Martin, A., Nadja, E. and Thomas, S.
REFERENCE	Nucleic acid encoding insect acetyl choline receptor subunit
AUTHORS	Patent: JP 2000023680-A 3 25-JAN-2000;
TITLE	
JOURNAL	HAVER AG

```

OS Heliothis virescens
EN JP 2000023680-A/3
PD 25-JAN-2000
PF 26-APR-1999 JP 1999118159
PR 04-MAY-1998 DE 19819829.9
PI MARTIN ADAMUTSUUSUKI,NADJA ERASU,THOMAS SCHULTE PC
C12N15/09,A01K67/033,C07K14/705,C07K16/28,C12N1/21,C12N5/10,PC
C12Q1/68,
PC G01N33/15,G01N33/50// (C12N1/21,C12R1:19),C12N15/00,C12N5/00 CC

FH Key Location/Qualifiers
FT CDS Location/Qualifiers
    I. 3109
    /organism="Heliothis virescens"
    /mol_type="genomic DNA"
    /db_xref="taxon:7102"

```

Qy	18	-----LeuLeuMetCysLeuLeuTrpProArgGlyAlaArgCysGly	31
Db	98	GCCTCATGTGTCGGCGCTGCTGGCTTTTCTGCTGCCGTATCGGAGCAA--GGT	154
Qy	32	TyrHisGluLeuArgLeuLeuHisHisLeuLeuAspHisTyrAsnValLeuGluArgPro	51
Db	155	CCTCAGCAGAAGAGATCTCTGAACCGGTGCTGGCAACTTACAAACCCCTGGAGCCACCG	214
Qy	52	ValValAsnGluSerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIle	71
Db	215	GTGCCCAACGAGAGCGAACCCTAGAGGTTCAGGTCGGCTTGACCTTGCAGCAAAATCATT	274
Qy	72	AspValAspGluLysAsnGlnLeuLeuThrAsnIleTrpLeuLysLeuGluTrpAsn	91
Db	275	GAGCTGGACGAGAAGAAATCAACTACTATATAACCAATATATGGCTGCTGTTGGAGTGGAAAT	334
Qy	92	AspMetAsnLeuArgTrpAsnThrSerAspPheGlyGlyValLysAspLeuArgValPro	111
Db	335	GACTACAACTGAGGTGGAAACGACGAGTAGTATGGCGGGTCAAGGACCTCAGGATCAGC	394
Qy	112	ProHisArgLeuTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAsp	131
Db	395	CCCAACAAGTTGTGGAAAGCGGACGTCCTATGTATTAATAGTGTCTCAGCAGGGGTTTGAC	454
Qy	132	SerThrTyrProThrAsnValValArgAsnAsnGlySerCysLeuTyrValProPro	151
Db	455	GGGACCTACAGACCAACAGTGGTGGTCAGAGCGCGCGAGCTGTCCTGTACGTGCCACCT	514
Qy	152	GlyIlePheLysSerThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArg	171
Db	515	GGCATATTCAAGAGACATGCAAGATGGGACATCGCGTGGTTTCCCTTCGACGACCAACAC	574
Qy	172	CysGluMetLysPheGlySerTrpThrTyrAspGlyTyrGlnLeuAspLeuGlnLeuGln	191
Db	575	TGTGATATGAAGTTCGGTAGCTGGACATATAGCGGCAATCAGTTGGATCTGGTGCTATAAA	634
Qy	192	AspGluGlyGlyGlyAspIleSerSerPheValThrAsnGlyGluTrpGluLeuIleGly	211
Db	635	GATGAGCGAGCGGCGATCTATCGGACTTCATAAACAATGGGAGTGGTATCTAATAGGA	694
Qy	212	ValProGlyLysArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIle	231
Db	695	ATGCCAGGCAAAAGAACAAATAACATACGGGTGCTGCCCGAGCGCTTACGFTGGACGTC	754
Qy	232	ThrPheAlaValValIleArgLysThrLeuTyrTyrPhePheAsnLeuIleValPro	251
Db	755	ACCTTCACCATCATGATAGAGACGAACTTGTACTACTTCTTCAACCTGATCGTCCCG	814
Qy	252	CysValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGlu	271
Db	815	TGGCTGCTGATCTCATCGATGGCACCTCTCGGCTTCACACTGCGCACCACTCCGGAGAG	874
Qy	272	LysLeuSerLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAla	291
Db	875	AAACTCACACTTGGAGTCACTATTCTTCTATCGCTACCGGTGTTCTCTCAACCTGGTAGCC	934
Qy	292	GluThrMetProAlaThrSerAspAlaValProLeuLeuGlyThrTyrPheAsnCysIle	311
Db	935	GAGACCTTGCACAGGCTCTCCGACGCTATCCCTCTGTAGGACGTTACTTCAATTCATC	994
Qy	312	MetPheMetValAlaSerSerValValSerThrIleLeuIleLeuAsnTyrHisHisArg	331
Db	995	ATGTTTCATGGTAGCGTCTGTGGTACGTACGACGTGGTGGTACTCAATTACACCATCGA	1054
Qy	332	HisAlaAspThrHisGluMetSerAspTrpIleArgCysValPheLeuTyrTrpLeuPro	351
Db	1055	ACAGCTGATATACATGAAATGCCACAGTGGATAAAATCAGTATTCTTCAATGTGTGCCA	1114
Qy	352	TrpValLeuArgMetSerArgProGlySerAlaThrThrProProAlaArgVal---	370
Db	1115	TGGTACTCGAATGTCTGAGGCCAGGAGAGATCACCAGGAGAGACTATAATGATGTAAC	1174
Qy	371	ProProProAspLeuGluLeuArgGluArgTetSerSerLysSerLeuLeuAlaAsnVal	390

```
Db 1175 ACAGGATGAGGAGCTGGAAGTGAAGAGAGTCTGCGAAGTCTTGTCTGGCGAATGTT 1234
Qy 391 LeuAspIleAspAspPheArgHis-----ProGlnAlaGlnGln----- 404
Db 1235 CTAGATATTGATGATGACTTCAGACACAGCGCCCTCCGCTCCTAAACAGTACTGCTCGACC 1294
Qy 405 -----ProGlnCysCysArgTyr-----TyrArgGlyGly----- 414
Db 1295 GGGAAITTTGGACCTGGGTGCTCAATATTCGCGACGGATTTCCGTGGGTGCTGTCGGT 1354
Qy 415 -----GluGluAsnGlyAlaGlyLeuAlaAlaHisSerCysPheGlyValAsp 430
Db 1355 CCGTCCAGATGGAAGAGCTGGCGGGCTGGGTAGCCAC-----CAT 1399
Qy 431 TyrGluLeuSerLeuIleLeuLysGluIleArgValIleThrAspGlnMetArgLysAsp 450
Db 1400 CGCGAGCTGCACCTCATATGAGAGAGCTGCAGTTCATCAGCGCAGGATGAAGAAGGCT 1459
Qy 451 AspGluAspAlaAspIleSerArgAspTrpLysPheAlaAlaMetValValAspArgLeu 470
Db 1460 GATGAGAAGCCGAGCTGATCAGCAGCTGGAGTTTGTCTGGATGGTTGTTGATAGGTTT 1519
Qy 471 CysLeuIleIlePheThrLeuPheThrIleIleAlaThrLeuAlaValLeuSerAla 490
Db 1520 TGCCTGTTGCGTTCACACTTTTCACAATCATCGGACAGTAGTGTCTCTGTTATCGCA 1579
Qy 491 ProHisIleMetVal 495
Db 1580 CCGCATATCATCGTG 1594

RESULT 11
AX009614
LOCUS AX009614 3109 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 5 from Patent EP0962528.
ACCESSION AX009614
VERSION AX009614.1 GI:9996846
KEYWORDS
SOURCE
ORGANISM Heliothis virescens (tobacco budworm)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuoidea; Noctuidae; Heliothinae; Heliothis.
REFERENCE 1
Adamczewski M.D., Schulte T.D. and Oellers, N.D.
Nucleic acids encoding acetylcholin-receptor subunits from insects
Patent: EP 0962528-A 5 08-DEC-1999;
BAYER AG (DE)
FEATURES
source
1..3109
location/Qualifiers
1..3109
/organism="Heliothis virescens"
/mol_type="unassigned DNA"
/db_xref="taxon:7102"
95..1600
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC07501.1"
/db_xref="GI:9996847"
/db_xref="REFMREBL:CAC07501"
/translation="MAPMLAALALLPVSEQGPHEKRLINLANKYNTLERPVANE
SEPLEVFLGLTQIIDVDEKNOLLITNIWLSLEWYDNLNWDSEYGVKDLRTPN
KLWKPDLVMSNADGPDGTQTNVNVRSGLYVPPGPKSTCKMDIAWFPDDOH
CDMKFGSWTDYGNQLDLVADAGGLSDPTTNGEWYLIIGMPKKNITVYACPEPV
DVFTTIRMRTLYFFNLIVPCVLISSMALLGFTLPDPSGEKLTIGVTLISLTIVFL
NLVAETLPQVSDAIPLLGTFTNCFIMFVASSVLTVVLNHYHRTADIEHMPQWIKSV
FLOWPLWLRSPGKKITRKNIMNTRMRELSKSLANLVIDIDDFRHGPP
PPNSTASTGNLPGPCSTFRIDFRSFVRPSTMEDVGGGLGSHREHLILIRELFITA
RMKKADEABELISDWKFAAMVWDRFCLFVTLFIITAVVLISAPHIIIVQ"
ORIGIN
Alignment Scores:
Pred. No.: 2.4e-164 Length: 3109
Score: 1809.00 Matches: 352
```

```
Percent Similarity: 77.52% Conservative: 55
Best Local Similarity: 67.05% Mismatches: 78
Query Match: 67.88% Indels: 40
DB: 6 Gaps: 8
US-09-303-232-4 (1-496) x AX009614 (1-3109)
Qy 5 AlaArgArgSerHisLeuAlaAlaProAlaGlyLeuLeu----- 17
Db 38 GCAGCGCCGCGTGTAGCGCAGCGTGTGCGCGGTGTGTATGTCGCGCGTGTGCGCGCTATG 97
Qy 18 -----LeuLeuLeuCysLeuLeuTrpProArgGlyAlaArgCysGly 31
Db 98 GCCCTATGTTGGCGGCTTGGCGCTGCTGGCTTGTGCCCCGTATCGGACAA---GGT 154
Qy 32 TyrHisGluLysArgLeuLeuHisLeuLeuAspHisTyrAsnValLeuGluArgPro 51
Db 155 CCTCAGAGAAGAGACTCTTGAACGCGTTGCTGCGAACTACAACACCCCTGGAGCGACCG 214
Qy 52 ValValAsnGluSerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleIle 71
Db 215 GTGCCCAACGAGAGCGAACCCGCTAGAGGTCAGGTTTCGGCTTGACCTTGCAACAAATCAT 274
Qy 72 AspValAspGluLysAsnGlnLeuLeuIleThrAsnIleTrpLeuLysLeuGluTrpAsn 91
Db 275 GACGTGACGAGAAGAATCAACTACTTATTAACCAATATATGCTGTGTTGGAGTGAAT 334
Qy 92 AspMetAsnLeuArgTrpAsnThrSerAspPheGlyGlyValLysAspLeuArgValPro 111
Db 335 GACTACAACTGAGGTGGAAACGACAGCGAGTATGCGGGGTCAAGGACCTCAGGATCAGC 394
Qy 112 ProHisArgLeuTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAsp 131
Db 395 CCCAACAAAGTTGTGGAAAGCGGACGCTCTTATGATATAATAGTGTGACGAGGGTTTGAC 454
Qy 132 SerThrTyrProThrAsnValValArgAsnGlnSerCysLeuTyrValProPro 151
Db 455 GGGACCTACAGACCAACGCTGTGTGTCAGAAAGCGGCGGAGTGCCTGTAGTCCACCT 514
Qy 152 GlyIlePheLysSerThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArg 171
Db 515 GGCATATTCAAGAGACACATCAAGATGGACATCGCGTGGTTTTCCTTCGACGACCAACAC 574
Qy 172 CysGluMetLysPheGlySerTrpThrTyrAspGlyTyrGlnLeuAspLeuGlnLeuGln 191
Db 575 TGTGATATGAAGTTTCGGTAGCTGGACATATGACGGCAATCAGTTGGATCTCTGGTCTAAA 634
Qy 192 AspGluGlyGlyAspIleSerSerPheValThrAsnGlyGluTrpGluLeuIleGly 211
Db 635 GATGAGCGAGCGGCGATCTATCGGACTTCATACAAATGGGAGTGGTATCTAATAGGA 694
Qy 212 ValProGlyLysArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIle 231
Db 695 ATGCGAGCAAAAAGAACACATAACATACGCGTGTGTCGCCCGGAGCCCTACGTGGACGTC 754
Qy 232 ThrPheAlaValIleAlaArgArgLysThrLeuTyrTyrPhePheAsnLeuIleValPro 251
Db 755 ACCTTACCACATCATGATGAAGAAGCAACCTTGTACTACTTCTTCAACCTGATCGTCCCG 814
Qy 252 CysValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGly 271
Db 815 TGGGTGCTGATCTCATCGATGGCACTCTCTCGGCTTCACATGCCACCACTCCGGAGAG 874
Qy 272 LysLeuSerLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAla 291
Db 875 AAACCTCACACTTGGAGTCACTATTCTTCTATCGCTGACGCGTGTCTCTCAACCTGGTAGCC 934
Qy 292 GluThrMetProAlaThrSerAspAlaValProLeuLeuGlyThrTyrPheAsnCysIle 311
Db 935 GAGACCCCTGCCAGGCTCTCCGACGCTATCCCGCTGTAGGAGCAGTACTTCAATTGTCATC 994
Qy 312 MetPheMetValAlaSerSerValSerThrIleLeuIleLeuAsnTyrHisHisArg 331
```


445 GAAAGCTGTCAAGGACCTCATGAAAAAGCGCTGCTGAACCATCTGCTGTCCACCTACAAT 504
47 VallLeuGluArgProValValAsnGluSerAspProLeuGlnLeuSerPheGlyLeuThr 66
505 ACGTGTGAGGACCGGTGGCCAAATGAATCGAGAGCCCTCGAGGTTAAGTTTCGACTGACG 564
67 LeuMetGlnIleIleAspValAspGluLysAsnGlnLeuLeuIleThrAsnIleTrpLeu 86
565 CTGCAGCAGATCATCGATGGAGAGAGAAATCAGCTTCTCAATACAAATCTTTGGCTT 624
87 LysLeuGluTrpAsnAspMetAsnLeuArgTrpAsnThrSerAspPheGlyGlyValLys 106
625 TCGTTGGAGTGGAAACGACTACAATCTGCGCTGGAATGAAACGGAATACGGCGGGTCAAG 684
107 AspLeuArgValProProHisArgLeuTrpLysProAspValLeuMetTyrAsnSerAla 126
685 GATCTACGAATCAACGCCCAACAGCTGTGGAAAGCCCGACCGTCTCATGTACAACAGCGCG 744
127 AspGluGlyPheAspSerThrTyrProThrAsnValValArgAsnAsnGlySerCys 146
745 GATGAGGATTCGATGCGCATCACACCAACATTTGGTCAAAACATGGCGGCGATTGT 804
147 LeuTyrValProProGlyIlePheLysSerThrCysLysIleAspIleThrTrpPhePro 166
805 CTGTACGTGCCCTGGTATCTTCAAGAGCACATGCAAGATGGACATCACTGTTGCCA 864
167 PheAspGlnArgCysGluMetLysPheGlySerTrpThrTyrAspGlyTyrGlnLeu 186
865 TTTGATACCAACAATTCGAAATGAATTCGGTAGTTGGACTTACGATGGAATTCAGTTG 924
187 AspLeuGlnLeuGlnAspGluGlyGlyAspIleSerSerPheValThrAsnGlyGlu 206
925 GATTGGTTTGAATTCGAGATGGAGGGATCTTCCGATTTCAACAATATGGCGAG 984
207 TrpGluLeuIleGlyValProGlyLysArgAsnGluIleTyrTyrAsnCysCysProGlu 226
985 TGGTACTTGTGTCATGTCGGGAAAGAAATACGATAGTCTACGCTGCTGCCAGAA 1044
227 ProTyrIleAspIleThrPheAlaValIleArgArgLysThrLeuTyrTyrPhePhe 246
1045 CCATATGTCATATCACTTACTATACAAATTCGTCGCGTACATATATATTTTTC 1104
247 AsnLeuIleValProCysValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuPro 266
1105 AATTATATGTCGATGTGCTAATCTCATCGATGACCTACTGGCGTTACATGCGG 1164
267 ProAspSerGlyGluLysLeuSerLeuGlyValThrIleLeuLeuSerLeuThrValPhe 286
1165 CCGGATTCGGCGGAGAACTGACGCTGGCGGTAATCTACTATCATTAACAGTATTT 1224
287 LeuAsnMetValAlaGluThrMetProAlaThrSerAspAlaValProLeuLeuGlyThr 306
1225 CTAACCTTGTGCGGAGTCCATGCCACAAACGTCGGATGCTGTCTCTATAGGCACC 1284
307 TyrPheAsnCysIleMetPheMetValAlaSerSerValValSerThrIleLeuIleLeu 326
1285 TACTTCAATTCATCATGTTTCATGTCCTCGTGGTGGTCTGCACAGTAGTGGTCTC 1344
327 AsnTyrHisHisArgHisAlaAspThrHisGluMetSerAspTrpIleArgCysValPhe 346
1345 AACTACCACCATCGCACAGCGAGCATTCACGAGATGCCCGTGGATCAAGTCGGTTTC 1404
347 LeuTyrTrpLeuProTrpValLeuArgMetSerArgProGlySerAlaThrTrpProPro 366
1405 CTCAATGGCTGCGCTGGATCTTTCGAAATGGGTGACCGCGTGCAGAAATTACACGCAA 1464
367 ProAlaArgValPro---ProProProAspLeuGluLeuArgGlnArgSerSerLysSer 385
1465 ACAATACTATTAGCAATCCGATGAGAGCTGAGCTAAAGAGCGCTCTCCAAATCC 1524
386 LeuLeuAlaAsnValLeuAspIleAspAspPheArgHisProGlnAlaGlnGlnPro 405
1525 CTGCTGGCCAATGTCTCTCGACATCGACGACGACTTCGCGGCACACAATATCT----- 1575

Qy 406 GlnCysCysArgTyrTyrArgGlyGlyGluAsnGlyAlaGlyLeuAlaAala----- 423
Db 1576 -----GGCTCCAAACCGCCATTTGGTCTGCTCGCGCAGCTTC 1611
Qy 424 -----HisSerCysPheGlyValAspTyr---GluLeu 433
Db 1612 GGTGCGGCCCAACGCTGGAGGAGCATCACAGGCCCATCGGCTGCAATCAAAAGATCTT 1671
Qy 434 SerLeuIleLeuLysGluIleArgValIleThrAspGlnMetArgLysAspAspGluAsp 453
Db 1672 CATCTAATTTCTCAAGAAATTCGAATTTATACGCGCGGATGCCAAAGCTGACGACGAA 1731
Qy 454 AlaAspIleSerArgAspTrpLysPheAlaAlaMetValValAspArgLeuCysLeuIle 473
Db 1732 GCGGAATGTATCGCGCATTCGAAGTTTCGCGCAATGTTGTGGATAGATTGTTTAAT 1791
Qy 474 IlePheThrLeuPheThrIleIleAlaThrLeuAlaValLeuLeuSerAlaProHisIle 493
Db 1792 GTTTTCAGCTCTTCAGATTAITGCAACGGTTACGGTGCTGCTCTCGCTCCGCACATA 1851
Qy 494 MetVal 495
Db 1852 ATCGTG 1857
RESULT 13
AF321447
LOCUS
DEFINITION
Drosophila melanogaster nicotinic acetylcholine receptor Dalpha6 subunit variant type III (nAcRalpha-30D) mRNA, complete cds, alternatively spliced.
AF321447
AF321447.1 GI:20152848
Drosophila melanogaster (fruit fly)
ORGANISM
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
AUTHORS
TITLE
Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes, Dalphas, Dalpha6 and Dalpha7, in Drosophila melanogaster Identify a New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-Mediated A-to-I Pre-mRNA Editing
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
Direct Submission
JOURNAL
Genetics 160 (4), 1519-1533 (2002)
11973307
2 (bases 1 to 2023)
Grauso, M. and Sattelle, D.B.
Submitted (15-NOV-2000) Human Anatomy and Genetics-Functional Genetics Unit, University of Oxford, South Park Road, Oxford OX1 3QX, UK
FEATURES
source
1. .2023
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon.7227"
/chromosome="2"
/map="30D1"
/dev_stage="embryo"
1. .2023
/gene="nAcRalpha-30D"
379. .1863
/gene="nAcRalpha-30D"
/note="ion channel; neurotransmitter transmembrane receptor; alternatively spliced; contains exons 3b and 8b"
/codon_start=1
/product="nicotinic acetylcholine receptor Dalpha6 subunit variant type III"
/protein_id="AA013394.1"
/db_xref="GI:20152849"

SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephyaridae; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1665)
AUTHORS Millar, N.S.
TITLE Direct Submission
JOURNAL Submitted (31-MAR-2003) Millar N.S., Department of Pharmacology,
University College London, Gower Street, London, WC1E 6BT, UNITED KINGDOM

FEATURES
source Location/Qualifiers
1..1665 /organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
gene 1..1665
/gene="nAcRalpha-30D"
CDS 44..1528
/gene="nAcRalpha-30D"
/codon_start=1
/product="nicotinic acetylcholine receptor subunit Dalphaf"

sig_peptide 44..121
mat_peptide 122..1525
/gene="nAcRalpha-30D"
/product="nicotinic acetylcholine receptor subunit Dalphaf"

ORIGIN
Alignment Scores:
Pred. No.: 4,84e-163 Length: 1665
Score: 1791.50 Matches: 342
Percent Similarity: 78.88% Conservative: 54
Best Local Similarity: 68.13% Mismatches: 81
Query Match: 67.22% Indels: 25
DB: 3 Gaps: 5

US-09-303-232-4 (1-496) x DME554209 (1-1665)

QY 8 SerHisLeuAlaProAlaGlyLeuLeuLeuLeuLeuCysLeuLeuTrpProArgGly 27
Db TCCCGCTGCCAGCGTCGTGCCTGGTGTTCCTGGATCAATTAA 109

QY 28 AlaArgCys---GlyPyrHisGluLysArgLeuLeuHisHisLeuLeuAspHisTyraSn 46
Db GAAAGCTGTCAGGACCTCATGAAGAAGCGCTGCTGAACCACTGCTCTCCACCTACAAT 169

QY 47 ValLeuGluArgProValValAsnGlnSerAspProLeuGlnLeuSerPheGlyLeuThr 66
Db ACGCTGGAGCGACCGTGCCCAATGAATCGGAGCCCTCGAGTTTAAGTTCGGACTGACG 229

QY 67 LeuMetGlnIleLeuAspValAspGluLysAsnGlnLeuLeuLeuLeuLeuLeuLeuTrpLeu 86
Db CTGCAGCAGATCATGCAGTGGACGAGGAATACAGTTCTCATAACCAATCTTTGGCTT 289

QY 87 LysLeuGluTrpAsnAspMetAsnLeuArgTrpAsnThrSerAspPheGlyGlyVallys 106
Db TCGTGGAGTGGNACGACTACAACTCTGGCTGCAATGAATGAACGGAATACGGCGGGTCAAG 349

QY 107 AspLeuArgValProProHisHisArgLeuTrpLysProAspValLeuMetTyraSnSerAla 126

350 GATCTACGAATCACGCCAACAACAGCTGTGGAAGCCCGAGCTGCTCATGTACACACGCGG 409
QY 127 AspGluGlyPheAspSerThrTyrrProThrAsnValValValArgAsnAsnGlySerCys 146
Db GATGAGGATTCGATGCGCATCATCACACGAGCTTGTGTCAAACATGGCGCAGTTGT 469

QY 147 LeuTyrrValProProGluGlyPheLysSerThrCysLysLysLysLysLysLysLysLys 166
Db CTGTACGTGCCCCCTGTGTTCATTCAAGAGCATGCAAGATGACATCAGTGGTGTCCCA 529

QY 167 PheAspAspGlnArgCysGluMetLysPheGlySerThrTyrrTyrrAspGlyTyrrGlnLeu 186
Db TTGTGATGCCAACATTCGCAAAATGAAATTCGGTAGTTGGACTTACGATGGAATCAGTTG 589

QY 187 AspLeuGlnLeuGlnAspGluGlyGlyAspLysSerSerPheValThrAsnGlyGlu 206
Db GATTTGGTTTTGAGTTCCGAAGATGGAGGGGATCTTCGGATTTCATAACAAATGGCGAG 649

QY 207 TrpGluLeuIleGlyValProGlyLysArgAsnGluIleTyrrTyraSnCysCysProGlu 226
Db TGGTACTTGTCTCCATGCGGGAAGAAGAAATACGATAGTCTAGCCCTGCTGCCCGACAA 709

QY 227 ProTyrrIleAspIleThrPheAlaValValIleIleArgArgLysThrLeuTyrrPhePhe 246
Db CCATATGTCGATATCACCTTACTATACAAATTCGTGCGCGTACATATATATATTTTTC 769

QY 247 AsnLeuIleValProCysValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuPro 266
Db AAATTAATTTGTCATGTGTCTAATCTCATCGATGGCGCTACTGCGGCTTCAATGCGG 829

QY 267 ProAspSerGlyGluLysLeuSerLeuGlyValThrIleLeuLeuSerLeuThrValPhe 286
Db CCGGATTCGGCGAGAAACTGACCTGGCGGTPAATCTACTACTACTATCATACATATTT 889

QY 287 LeuAsnMetValAlaGluThrMetProAlaThrSerAspAlaValProLeuLeuGlyThr 306
Db CTAAACCTTGTGCGGAGTCCATGCGCACACGTCGGATGCTGTTCTCTTATAGCACCC 949

QY 307 TyrPheAsnCysIleMetPheMetValAlaSerSerValValSerThrIleLeuLeuLeu 326
Db TACTTCAATTTGTCATGTTCATGTGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1009

QY 327 AsnTyrrHisArgHisAlaAspThrHisGluMetSerAspTrpIleArgCysValPhe 346
Db AACCTACACATCGCACAGCGGACATTCACGAGATGCCACCGTGGATCAAGTCCGTTTTC 1069

QY 347 LeuTyrrTrpLeuProTrpValLeuArgMetSerArgProGlySerAlaThrTrpPro 366
Db CTACAATGCTGCTCGTGGATCTTCGAAATGGCTGACCGGTCGCAAGATTACACGCAA 1129

QY 367 ProAlaArgValPro---ProProAspLeuGluLeuArgGluArgSerSerLysSer 385
Db ACAATACTATTAAAGCAATCGCATGAAGGAGCTGGAGCTAAAAGAGCGCTCTCCAAATCC 1189

QY 386 LeuLeuAlaAsnValLeuAspIleAspAspPheArgHisProGlnAlaGlnGlnPro 405
Db CTGCTGGCCAATGCTCTGACATCGACACGACTTCCGGGACACAAATATCT----- 1240

QY 406 GlnCysCysArgTyrrArgGlyGlyGluGluAsnGlyLaGlyLeuAlaAala----- 423
Db -----GGTCCCACAAACCGCCATTGGCTGCTGCGCCGAGCTTC 1276

QY 424 -----HisSerCysPheGlyValAspTyr---GluLeu 433
Db GGTGCGCCCAACACGGTGGAGGAGCATCACACGGCCATCGGCTGCAATCAAAAGATCTT 1336

QY 434 SerLeuIleLeuLysGluIleArgValIleThrAspGlnMetArgLysAspAspGluAsp 453
Db CATCTAATTTCTCAAGAATTTGCAATTTATTACCGCGCGGATGCGCAAGCTGACGACGAA 1396

QY 454 AlaAspIleSerArgAspTrpLysPheAlaMetValValAspArgLeuCysLeuIle 473

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 8, 2004, 06:39:04 ; Search time 520.982 Seconds

(without alignments)
4044.486 Million cell updates/sec

Title: US-09-303-232-4

Perfect score: 2665

Sequence: 1 MCGRRARRSHLAAPAGLLLLL.....LFTIATLAVLLSAPHIMVS 496

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US09303232/runat_07052004_101109_23873/app_query.fasta_1.2261
-DB=N_Geneseq_29Jan04 -QWIT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09303232_@CGN_1_1200_@runat_07052004_101109_23873 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N Geneseq_29Jan04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001as:*
5: Geneseq2001bs:*
6: Geneseq2002s:*
7: Geneseq2003as:*
8: Geneseq2003bs:*
9: Geneseq2003cs:*
10: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2665	100.0	3700	3 AAZ24476	Aaz24476 H. viresc
2	1850	69.4	2886	3 AAZ24475	Aaz24475 D. melano
3	1809	67.9	3109	3 AAZ24477	Aaz24477 H. viresc
4	1623.5	60.9	1540	4 ABL07231	AbL07231 Drosophil
5	1417.5	53.2	936	4 ABL13733	AbL13733 Drosophil
6	1267	47.5	2769	2 AAT59196	Aat59196 Human neu
7	1247	46.8	1876	2 AAV12197	Aav12197 Human neu
8	1247	46.8	1876	2 AAT48239	Aat48239 Neuronal

9	1247	46.8	1876	6 ABS54875	AbS54875 Human neu
10	1247	46.8	1876	6 ABV73248	Abv73248 Human neu
11	1247	46.8	1876	8 ADA10864	Ada10864 Human neu
12	1246	46.8	1509	3 AAC58195	Aac58195 Human PRO
13	1246	46.8	1509	4 AAC90380	Aac90380 Wild-type
14	1242	46.6	1590	2 AAV44687	Aav44687 V2747 var
15	1240	46.5	1509	4 AAC90385	Aac90385 Mutant hu
16	1236	46.4	1509	4 AAC90386	Aac90386 Mutant hu
17	1230	46.2	1509	4 AAC90387	Aac90387 Mutant hu
18	1219.5	45.8	2101	2 AAT59197	Aat59197 Neuronal
19	1215.5	45.6	1964	6 ABZ11298	Abz11298 Human pol
20	1193	44.8	803	4 ABL07799	AbL07799 Drosophil
21	1057.5	39.7	1915	4 AAD20962	Aad20962 Caenorhab
22	1011	37.9	5096	4 ABL13732	AbL13732 Drosophil
23	1003.5	37.7	1416	4 AAC90382	Aac90382 ChimERIC
24	968	36.3	1896	6 AAL45867	Aal45867 Modified
25	960	36.0	1908	2 AAT48236	Aat48236 Neuronal
26	960	36.0	1908	8 ADA10856	Ada10856 Human neu
27	960	36.0	2277	2 AAV12199	Aav12199 Human neu
28	960	36.0	2277	2 AAQ90387	Aaq90387 Alpha 2 s
29	960	36.0	2277	6 ABS54870	AbS54870 Human neu
30	960	36.0	2277	6 ABV73243	Abv73243 Human neu
31	960	36.0	2664	2 AAT48235	Aat48235 Neuronal
32	960	36.0	2664	6 ABK92165	Abk92165 Prostate
33	960	36.0	2664	8 ADA10854	Ada10854 Human neu
34	960	36.0	2664	9 ADC71170	Adc71170 Human 205
35	960	36.0	2666	9 AAL45873	Aal45873 Modified
36	957.5	35.9	1869	6 ADA10862	Ada10862 Human neu
37	957.5	35.9	1869	6 ABL54794	AbL54794 Modified
38	954	35.8	1584	7 ADA83809	Ada83809 Human CHR
39	935.5	35.1	2210	4 ABL10275	AbL10275 Drosophil
40	929.5	34.9	1743	2 AAT48232	Aat48232 Neuronal
41	929.5	34.9	1743	6 ABS54874	AbS54874 Human neu
42	929.5	34.9	1743	6 ABV73247	Abv73247 Human neu
43	929.5	34.9	1743	8 ADA10862	Ada10862 Human neu
44	923.5	34.7	2385	4 ABL11821	AbL11821 Drosophil
45	913	34.3	1915	2 AAT48241	Aat48241 Neuronal

ALIGNMENTS

RESULT 1

AZ24476
ID AAZ24476 standard; CDNA to mRNA; 3700 BP.

XX AC AAZ24476;

DT 17-FEB-2000 (first entry)

DE H. virescens acetyl-choline receptor DNA from clone Hva7-1.

KW Acetyl-choline receptor; nicotinic; insect; insecticide; screening;
KW neurotransmission; plant protection agent; conductance; ACHR; ds.

OS Heliothis virescens.

FH Key Location/Qualifiers
CDS 335..1825

FT /*tag= a
FT /product= "acetyl-choline receptor"

PN D819819829-Al.

XX 11-NOV-1999.

XX 04-MAY-1998; 98DE-01019829.

XX 04-MAY-1998; 98DE-01019829.

XX (FARB) BAYER AG.

XX Adamczewski M, Oellers N, Schulte T;

DR WPI: 2000-014207/02.
 DR P-FSDB; AAY50815.
 XX New nucleic acid encoding a nicotinic acetylcholine receptor from
 PT insects, used to identify potential insecticides.
 XX Claim 1a; Page 14-17; 26pp; German.
 XX This invention describes a novel nucleic acid (NA) encoding a nicotinic
 CC acetylcholine receptor (I) from insects which can be used as an
 CC insecticide. Inhibitors of (I) interfere with neurotransmission. (I)
 CC (also vectors containing it, its regulatory regions, and antibodies
 CC directed against (I)-encoded proteins) are used to screen for: (a) plant
 CC protection agents that alter conductance of ACHR, potentially useful as
 CC insecticides, or (b) genes which encode polypeptides that are involved in
 CC formation of functionally related ACHR in insects. (I) are also used to
 CC isolate and characterize the specified regulatory regions and for
 CC recombinant production of (II). This sequence encodes an acetylcholine
 CC receptor isolated from *Heliothis virescens*
 XX
 SQ Sequence 3700 BP; 893 A; 953 C; 944 G; 910 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,72e-257 Length: 3700
 Score: 2665.00 Matches: 496
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0

US-09-303-232-4 (1-496) x AA224476 (1-3700)

QY 1 MetGlyArgAlaArgSerHisLeuAlaProAlaGlyLeuLeuLeu 20
 DB 335 ATGGCGGGCGGGCGCGCTCGACCTGGCGGGCGCGCGCTGCTGCTGCTG 394
 QY 21 CysLeuLeuTrpProArgGlyAlaArgCysGlyTyrHisGlyLeuLeuHis 40
 DB 395 TGCCCTGCTCTGCGCGGGGGGCGCGCTGCGGGTACCGAGAGCGGCTACTGGCAC 454
 QY 41 LeuLeuAspHisTyrAsnValLeuGluArgProValValAsnGluSerAspProLeu 60
 DB 455 CTATTGGACCACTCAACAGTACTGGAGAGCGCGCTCGTCAACGAGAGCGACCGCTGCG 514
 QY 61 LeuSerPheGlyLeuThreuleuMetGlnIleleAspValAspGluLeuLeu 80
 DB 515 CTCCTCTTGGCGCTCACGCTCATGAGATCATGAGCTGAGCGAGAGAACAGCTTTTA 574
 QY 81 IleThrAsnIleTyrLeuLysLeuGluTrpAsnAspMetAsnLeuArgTrpAsnThrSer 100
 DB 575 ATAAACAACATCTGGCTAAACTAGAGTGGATGATATGAATTTGAGGTGGACACTTCA 634
 QY 101 AspPheGlyValLysAspLeuArgValProProHisArgLeuTrpLysProAspVal 120
 DB 635 GATTTTCGGCGGGTCAAAGATTAAAGATGCGCACCCACAGACTATGGAACACAGACGTC 694
 QY 121 LeuMetTyrAsnSerAlaAspGluGlyPheAspSerThrTyrProThrAsnValVal 140
 DB 695 CTTATGTAAACAGCGCGGACGAAGGGTTGACAGACGATATCCCAACAGACCTGGTGG 754
 QY 141 ArgAsnAsnGlySerCysLeuTyrValProProGlyIlePheLysSerThrCysLysIle 160
 DB 755 CGGAACAACGGCTCGTGTCTGTACGTGCGCGCGGCACTTCAAGAGCACTTGAAGATC 814
 QY 161 AspIleThrTrpPheProPheAspAspGlnArgCysGluMetLysPheGlySerTrpThr 180
 DB 815 GACATCACCTGGTTCCTTCGACCAACAGATGCGAGATGAAGTTTGGCAGCTGGACT 874
 QY 181 TyrAsnGlyTyrGlnLeuAspLeuGlnLeuAspGluGlyGlyAspIleSerSer 200
 DB 875 TATGATGGTTATCAGTTGGATCTCAACCTACAGATGAAGGGGGGAGATATAGCAGT 934
 QY 201 PheValThrAsnGlyGluTrpGluLeuIleGlyValProGlyLysArgAsnGluIleTyr 220

DB 935 TTTGTACGAATGGGAATGGAGTTAATAGAGTCCCGCGCAAGCGCAACAGATCTAC 994
 QY 221 TyrAsnCysCysProGluProTyrIleAspIleThrPheAlaValValIleArgArgLys 240
 DB 995 TACAACGTGTTGTCGGAGCCATACATCGACATCAGTTTGGGTGGTATCCGGAGGAAA 1054
 QY 241 ThrLeuTyrTyrPhePheAsnLeuValProCysValLeuIleAlaSerMetAlaLeu 260
 DB 1055 ACGTCTACTACTTCTTCAATCTGATGTCCTTGGGTGCTCATCGCCTCCATGGCTCTA 1114
 QY 261 LeuGlyPheThrLeuProAspSerGlyGlyLysLeuSerLeuGlyValThrIleLeu 280
 DB 1115 TTGGGTTTCACTTGGCTCCAGACTCCGAGAGAAAAGTTGTCTTAGGTGTGACGATATTA 1174
 QY 281 LeuSerLeuThrValPheLeuAsnMetValIleGluThrMetProAlaThrSerAspAla 300
 DB 1175 CTGCTGTTGACGGTGTCTCAACATGGTGGCGAGACGATGCCAGACGCTCGACGCC 1234
 QY 301 ValProLeuLeuGlyThrTyrPheAsnCysIleMetPheMetValAlaSerSerValVal 320
 DB 1235 GTGCCCTGTGCGGACCTACTTCACTGCATCATGTTCTGTTGGCTTCTCTCCGCTGTC 1294
 QY 321 SerThrIleLeuIleLeuAsnTyrHisHisArgHisAlaAspThrHisGluMetSerAsp 340
 DB 1295 TCCACCATCTGATCTCAACTACCACCGCGCAGCAGACACTCACGAAATGAGTGAT 1354
 QY 341 TrpIleArgCysValPheLeuTyrTrpLeuProTrpValLeuArgMetSerArgProGly 360
 DB 1355 TGGATTCTGCTGCTTCTTATTGGCTGCGTGGGTGCTGCGCATGTCACGGCCCGGC 1414
 QY 361 SerAlaThrThrProProAlaArgValProProProAspLeuGluLeuArgGlu 380
 DB 1415 TCGGGACAGCGCGCGCGCGCGCTACTCTCGCGCGCGGACCTGGAGCTGCGCGAG 1474
 QY 381 ArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAspAspPheArgHisPro 400
 DB 1475 CGTCTCTCAAGTGTCTCTAGCGAAGCTGTCGACATCGATGACGACTTCCGCCACCG 1534
 QY 401 GlnAlaGlnGlnProGlnCysCysArgTyrTyrArgGlyGlyGluGluAsnGlyAlaGly 420
 DB 1535 CAAGCGCAGCAGCGCAATGCTCGCGATACACAGGGGGGGGAGAGAAATGCGCGGGG 1594
 QY 421 LeuAlaAlaHisSerCysPheGlyValAspTyrGluLeuSerLeuIleLeuLysGluIle 440
 DB 1595 TTGGCGGCGCAGCTTGTCTCGTGTGCTACTACGAGCTCTCCTCATTTCTGAAGGAGAT 1654
 QY 441 ArgValIleThrAspGlnMetArgLysAspAspGluAspAlaAspIleSerArgAspTrp 460
 DB 1655 AGAGTCATCACAGATCAGATGCGCAAGGACGACGAGATGCGGACATTTTCGCGCAGCTG 1714
 QY 461 LysPheAlaAlaMetValValAspArgLeuCysLeuIleIlePheThrLeuPheThrIle 480
 DB 1715 AAGTTCGCGCGCATGGTGTGAGACGACTGTGCTTATTTATTTTACCTGTTCACATC 1774
 QY 481 IleAlaThrLeuAlaValLeuSerAlaProHisIleMetValSer 496
 DB 1775 ATCGCACGCTAGCTGCTGCTGTCGCGCGCACACATCATGTTGCTG 1822
 RESULT 2
 AA224475
 ID AA224475 standard; cDNA to mRNA; 2886 BP.
 XX
 AC AA224475;
 XX
 DT 17-FEB-2000 (first entry)
 XX
 DE D. melanogaster acetylcholine receptor DNA from clone Da7.
 KW Acetylcholine receptor; nicotinic; insect; insecticide; screening;
 KW neurotransmission; plant protection agent; conductance; ACHR; ds.
 XX
 OS Drosophila melanogaster.


```

Db 890 ATTCTTAACCTTGTCCGCGAGTCCATGCGCACAGCTGGTGTCTCTTATAGG 949
Qy 305 yThr----- 306
Db 950 TAC-ACACACAGCATTTAAACAGCTATAGCTTAAAGTACGAGCCGATCGGTCCGACTAT 1008
Qy 307 -----TyrPheAsnCysIleMetPheMetVa 315
Db 1009 CGACCACTACTACCACTGTCGAGCAGAAAGCACCTACTTCAATTGCATCATGTTCAUGGT 1068
Qy 315 lalaSerSerValValSerThrIleLeuIleLeuAsnTyrHisHisArgHisAlaAspTh 335
Db 1069 CGCCTCGTGGTGTGTCAGAGTAGTGTCTCACTACCATCGCACAGCGGACAT 1128
Qy 335 rHisGluMetSerAspTrrIleArgCysValPheLeuTyrTrpLeuProTrpValLeuAr 355
Db 1129 TCACGAGATGCCACCGTGGATCAAGTCGCTTTTCTCAATGGCTGGCTGGATCTCGG 1188
Qy 355 gMetSerArgProGlySerAlaThrThrProProAlaArgValPro---ProProPr 374
Db 1189 AATGGTCCACCGGTCACAGATTACAGCAAAACAATACTATTAAAGCAATCGCATGAA 1248
Qy 374 oAspLeuGluLeuArgGluArgSerSerIysSerLeuLeuAlaAsnValLeuAspIleAs 394
Db 1249 GGAGCTGGAGCTAAAGGAGCGCTCTCCAAATCCCTGCTGGCCAATGTCTCGACATCGA 1308
Qy 394 pAspAspPheArgHisProGlnAlaGlnGlnProGlnCysCysArgTyrTyrArgGlyGl 414
Db 1309 CGACCACTTCGGGCACAAATATCT-----GG 1335
Qy 414 yGluGluAsnGlyAlaGlyLeuAlaAla----- 423
Db 1336 CTCCTCAACCGCATGCTGCTCGCGCAGCTTCGGTCGCCCAACACGCTGGAGGACGA 1395
Qy 424 -HisSerCysPheGlyValAspTyr---GluLeuSerLeuIleLeuLysGluIleArgVa 442
Db 1396 TCACACGGCCATCGGCTGCAATCACAAAGATCTTCTAATTTCTCAAAGAATTGCAATT 1455
Qy 442 lIleThrAspGlnMetArgLysAspAspGluAspAlaAspIleSerArgAspTrpLysPh 462
Db 1456 TATTACGGCGGAGTGGCGAAAGCTGACGACGAAGCAATTGATCGGCGATTGGAAAGTT 1515
Qy 462 eAlaAlaMetValValAspArg 469
Db 1516 CGCGGCAATGTTGTGGATAGG 1537

RESULT 5
ABL13733
ID ABL13733 standard; cdNA; 936 BP.
AC ABL13733;
XX
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 35681.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
XX
KW pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
FN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
PF 23-MAR-2001; 2000US-0191637P.
XX
PR 11-JUL-2000; 2000US-00614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX

```

```

PI Venter JC, Adams M, Li PWD, Myers EW;
XX WPI: 2001-656860/75.
DR P-FSDB; ABB69630.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 35681; 2lpp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
CC ABBS7072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 936 BP; 238 A; 213 C; 239 G; 246 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1.23e-132 Length: 936
Score: 1417.50 Matches: 265
Percent Similarity: 91.78% Conservative: 14
Best Local Similarity: 87.17% Mismatches: 16
Query Match: 53.19% Indels: 9
DB: 4 Gaps: 2
XX
US-09-303-232-4 (1-496) x ABL13733 (1-936)
Qy 2 GlyGlyArgAlaArgArgSerHisLeuAlaProAlaGlyLeuLeuLeuLeuLeuCys 21
Db 46 GGTGTCAGAAAGT-----CTGTCATATGCGCTGGGACTTTTATTATGATA --- 90
Qy 22 LeuLeuTrpProArgGlyAlaArgCysGlyTyrHisGluLysArgLeuLeuHisLeu 41
Db 91 -----CCGGCTTGTGGCGCTGACCCCATGAGAGCGCTACTCCACGCCCTT 138
Qy 42 LeuAspHisTyrAsnValLeuGluArgProValValAsnGluSerAspProLeuGlnLeu 61
Db 139 CTGGACAACTACACAGCCTGGAGCGTCGGTGGTCAATGAATCCGATCCATTCGCACTG 198
Qy 62 SerPheGlyLeuThrLeuMetGlnIleAspValAspGluLysAsnGlnLeuLeuIle 81
Db 199 AGCTTCGGACTAACACTCATGCAGATTATCGATGTGGACGAAAGAAATCACTGCTTATA 258
Qy 82 ThrAsnIleTrpLeuLysLeuGluTrpAsnAspMetAsnLeuArgTrpAsnThrSerAsp 101
Db 259 ACGAATATTGGCTCAAAATGGAAATGGAACGATATGATATGATCTTCGATGGAATTCGAGTGAG 318
Qy 102 PheGlyGlyValLysAspLeuArgValProProHisArgLeuTrpLysProAspValLeu 121
Db 319 TTCGTGGTGTGGGATCTCGCAATTCGCGACATCGCCTATGGAACCGGATGACTAG 378
Qy 122 MetTyrAsnSerAlaAspGluGlyPheAspSerThrTyrProThrAsnValValValArg 141
Db 379 ATGTACACAGTGGCGACGAGGGCTTCGATGGAACGTACGCCACAAATGTGGTGTTCGC 438
Qy 142 AsnAsnGlySerCysLeuTyrValProProGlyIlePheLysSerThrCysLysIleAsp 161
Db 439 AATAATGGAGCTGTCTGTACCTACCGCAGGTATATTATTAAGTCAACGCTGAAGATCGAC 498
Qy 162 IleThrTrpPheProPheAspAspGlnArgCysGluMetLysPheGlySerTrpThrTyr 181
Db 499 ATTACGTGGTTTCCATTTCGACATCAGATGTGAAATGAAATTTGGTTCTGGGACCTAC 558
Qy 182 AspGlyTyrGlnLeuAspLeuGlnLeuGlnAspGluGlyGlyAspIleSerSerPhe 201
Db 559 GATGGGTTTCAGTTGGACCTGTCAGTTGCAGGACGAAGCTGGTGGCGACATTTCCTAGCTTT 618

```


ID ABV73248 standard; cDNA; 1876 BP.
 XX
 AC ABV73248;
 XX
 DT 22-JAN-2003 (first entry)
 XX
 DE Human neuronal NACHR alpha7 subunit encoding cDNA.
 XX
 KW Human; neuronal; nicotinic acetylcholine receptor; NACHR; drug screening;
 KW immunochemistry; NACHR alpha7 subunit; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 73..1581
 FT /*tag= a
 FT /product= "NACHR alpha7 subunit"
 FT /note= "neuronal nicotinic acetylcholine receptor"
 XX
 PN W0200259266-A2.
 XX
 PD 01-AUG-2002.
 XX
 PF 29-OCT-2001; 2001WO-US050985.
 XX
 PR 01-NOV-2000; 2000US-00703951.
 XX
 PA (MERI) MERCK & CO INC.
 XX
 PI Gillespie A, Claeps BO, Chavez-Noriega LE, Siegel R, Elliott KJ;
 XX
 DR WPI; 2002-698532/75.
 DR P-PSDB; ABB82435.
 XX
 XX Cell comprising nucleic acids encoding human alpha and beta subunits of
 PT neuronal nicotinic acetylcholine receptors, useful for in vitro screening
 PT of a drug substance in a test system specific for humans.
 XX
 PS Example; Page 129-130; 143pp; English.
 XX
 CC The invention relates to a suitable host cell transfected with an
 CC isolated nucleic acid molecule comprising a sequence of nucleotides or
 CC ribonucleotides that encodes at least one alpha or beta subunit of a
 CC human neuronal nicotinic acetylcholine receptor (NACHR). The compositions
 CC and methods of the present invention, which provide a means to prepare
 CC synthetic or recombinant receptors and receptor subunits that are
 CC substantially free of contamination from many other receptor proteins,
 CC are useful for observing the effect of a drug substance on a particular
 CC subtype to perform in vitro screening of the drug substance in a test
 CC system that is specific for humans. The antibodies can be used in
 CC immunochemistry and for diagnostic and therapeutic applications. The
 CC present sequence represents a human neuronal NACHR alpha7 subunit
 CC encoding cDNA
 XX
 SQ Sequence 1876 BP; 369 A; 553 C; 530 G; 423 T; 0 U; 1 Other;
 Alignment Scores:
 Pred. No.: 4,91e-115 Length: 1876
 Score: 1247.00 Matches: 247
 Percent Similarity: 60.47% Conservative: 85
 Best Local Similarity: 44.99% Mismatches: 129
 Query Match: 46.79% Indels: 88
 DB: 6 Gaps: 8
 US-09-303-232-4 (1-496) x ABV73248 (1-1876)
 QY 3 GlyArgAlaArg-----SerHisLeuAlaAla--- 12
 XX
 DB 1 GCGCGAGCGGCGAGCGCGGCGACAGCGAGAGCGTGGAGCGCGCGCGCTCGTGCAGCT 60
 QY 13 -----ProAlaGlyLeuLeuLeuLeuLeuLeuLeuT 24
 XX
 DB 61 CCGGGACTCAACATGCGTCTCGCGGAGGCGTCTGGCTGCGCGCGTGGCTG 120

QY 25 ProArgGlyAlaArgCysGlyTyrHisGluLysArgLeuHisHisLeuLeuAspHis 44
 DB 121 CTGCACGTGCTCCCTGCAAGCGAGTTCACAGGAAGCTTTACAAGGAGCTGGTCAAGAAC 180
 QY 45 TyrAsnValLeuGluArgProValValAsnGluSerAspProLeuGlnLeuSerPheGly 64
 DB 181 TACAATCCCTTGGAGAGCGCGTGGCCAATGATCGCAACCACTCACCGCTACTCTCC 240
 QY 65 LeuThrLeuMetGlnIleAspValAspGluLysAsnGlnLeuLeuIleThrAsnIle 84
 DB 241 CTGAGCTCTGTCAGATCATGGATGAGAGAACCAAGTATTACCAACCAACATT 300
 QY 85 TrpLeuLysLeuGluTrpAsnAspMetAsnLeuArgTrpAsnThrSerAspPheGly 104
 DB 301 TGGCTGCAATGCTTGGACAGATCACTATTATACAGTGAATGTGTGAGAATATCCAGG 360
 QY 105 ValLysAspLeuArgValProHisArgLeuTrpLysProAspValLeuMetTyrAsn 124
 DB 361 GTGAAGACTGTTCTGTTTCCCAAGATGGCAGATTGGAAACAGACATTCTTCTATAAC 420
 QY 125 SerAlaAspGluGlyPheAspSerThrTyrProThrAsnValValArgAsnAsnGly 144
 DB 421 AGTGCTGATGAGCGCTTGGACGACATTCACACTACGTGTGGTGAATCTCTCTGG 480
 QY 145 SerCysLeuTyrValProGlyIlePheLysSerThrCysLysIleAspIleThrTrp 164
 DB 481 CATTCGCCAGTACCTCCCTCCAGGCATATTCAAGAGTTCCTGCTACATCATGATGACGCTGG 540
 QY 165 PheProPheAspAspGlnArgCysGluMetLysPheGlySerThrTrpTyrAspGlyTyr 184
 DB 541 TTTCCCTTTGATGTGCAGCACTGCAAACTGAAGTTGGTCTCTGCTGTAGCGAGCTGG 600
 QY 185 GlnLeuAspLeuGlnLeuAspGluGlyGlyAspIleSerSerPheValThrAsn 204
 DB 601 TCCTTGGATCTGAGATGACGAG-----GCAGATATACGTGGCTATATCCCCAAT 651
 QY 205 GlyGluTrpGluLeuIleGlyValProGlyLysArgAsnGluIleTyrTyrAsnCysCys 224
 DB 652 GGAGATGGGACCTAGTGGGAATCCCGGCAAGAGGAGTGAAGGTTCTATGAGTGTCTG 711
 QY 225 ProGluProTyrIleAspIleThrPheAlaValValIleArgArgLysThrIleuTyr 244
 DB 712 AAAGAGCCCTACCCCGATGTCACTTCACGTGACCATGCGCGCGGAGGACGCTACATAT 771
 QY 245 PhePheAsnLeuIleValProCysValLeuIleAlaSerMetAlaLeuLeuGlyPheThr 264
 DB 772 GGCCTCAACCTGCTGATCCCTGTGTGTCTATCTCGGCTCGCCTGCTGTGTCTCTG 831
 QY 265 LeuProProAspSerGlyLysLeuSerLeuGlyValThrIleLeuLeuSerIleuThr 284
 DB 832 CTTCTGCGAGATTCCGGGGAGAAGATTCCCTGGGGATAACAGTCTTACTCTCTTACC 891
 QY 285 ValPheLeuAsnMetValAlaGluThrMetProAlaThrSerAspAlaValProLeuLeu 304
 DB 892 GTCTTCATGCTGCTGCTGAGATCATGCGCGCAACATCCGATTCGATGATGATGATA 951
 QY 305 GlyThrTyrPheAsnCysIleMetPheMetValAlaSerSerValValSerThrIleLeu 324
 DB 952 GCCAGTACTTCGCCAGCACCATGATCATCGTGGGCTCTCGTGTGTGTGACGTGATC 1011
 QY 325 IleLeuAsnTyrHisArgHisAlaAspThrHisGluMetSerAspTrpIleArgCys 344
 DB 1012 GTGCTGCAGTACCAACCAACCGACCGCGGGGCAAGATCCCAAGTGGACCAAGTCT 1071
 QY 345 ValPheLeuTyrTrpLeuProTrpValLeuArgMetSerArgProGly----- 360
 DB 1072 ATCTTCTGAATGTTGCGCGTGGTCTCTSCGATGAAGAGGCGCGCGGAGGACAGGTC 1131
 QY 361 -----SerAla 362
 DB 1132 CGCCCGGCTGCCAGCAACAGCAGCGGCGCTGCAGCGCTGGCAGTGTGGAGATGACGCGC 1191

Percent Similarity:	62.45%	Conservative:	86
Best Local Similarity:	45.98%	Mismatches:	12
Query Match:	46.75%	Indels:	70
DB:	3	Gaps:	6

US-09-303-232-4 (1-496) x AAC58395 (1-1509)

QY 12 AlaProAlaGlyLeuLeuLeuLeuCysLeuLeuTrpProArgGlyAlaArgCysGly 31
Db :::: ||| :::: |||
10 TCGCGGGAGCGCTGTGGCTGGCGCTGCCTCTGACGTCCTCTGCAAGC 69
QY 32 TyrHisGluLysArgLeuLeuHisHisLeuLeuAspHisTyrAsnValLeuGluArgPro 51
Db :::: ||| :::: |||
70 GAGTTCCAGAGNAAGCTTTACAAAGAGCTGGTCAAGAATACAATCCCTTGAGAGGCC 129
QY 52 ValValAsnGlnSerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleLeu 71
Db :::: ||| :::: |||
130 GTGGCCAATGAAGCTCGCAACCACTCACCGCTACTTCTCCCTGAGCCTCTCGCATCATG 189
QY 72 AspValAspGluLysAsnGlnLeuLeuIleThrAsnIleTrpLeuLysLeuGluTrpAsn 91
Db :::: ||| :::: |||
190 GACGTGGATGAGAGAACCAAGTTTTAACCACCAACATTGGCTGCAATGTCTGGACA 249
QY 92 AspMetAsnLeuArgTrpAsnThrSerAspPheGlyGlyValLysAspLeuArgValPro 111
Db :||| ::||| :::: |||
250 GATCACTATTACAGTGAATGTGCAGAAATATCCAGGGGTGAAGACTGTTGGTTCCCA 309
QY 112 ProHisArgLeuTriplysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAsp 131
Db :::: ||| :::: |||
310 GATGCCAGATTGGAAACCAACACATCTCTCTATAACAGTGCCTGATGAGCGCTTGAC 369
QY 132 SerThrTyrProThrAsnValValArgAsnAsnGlySerCysLeuTyrValProPro 151
Db :||| ::||| :::: |||
370 GCCACATTCACACATAACGTGTGTGTAATCTCTGGGCATTGCCGACTCTGCTCTCA 429
QY 152 GlyIlePheLysSerThrCysLysIleAspIleThrTrpPheProPheAspGlnArg 171
Db :||| ::||| :::: |||
430 GGCAATATCAAGAGTTCCTGTCTACATCATGATGATCGCTGTCTGCTTTGATGTGCAGC 489
QY 172 CysGluMetLysPheGlySerTrpThrTyrAspGlyTyrGlnLeuAspLeuGlnLeuGln 191
Db :||| ::||| :::: |||
490 TGCAAACCTGAAGTTGGCTCTGGTCTTACGGAGCTGGTCTTGATCTGCAGATGAGC 549
QY 192 AspGluGlyGlyGlyAspIleSerSerPheValThrAsnGlyGluTrpGluLeuGly 211
Db :::: ||| :::: |||
550 GAG-----GCAGATATCAGTGGCTATATCCCCAATGAGAAATGGACCTAGTGGGA 600
QY 212 ValProGlyLysArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIle 231
Db :||| ::||| :::: |||
601 ATCCCCGGCAGAGAGTGAAGGTTCTATGAGTGTCTCAAGAGCCCTACCCGATGTC 660
QY 232 ThrPheAlaValValIleArgAllysThrLeuTyrTyrPhePheAsnLeuIleValPro 251
Db :||| ::||| :::: |||
661 ACCTTCAAGTAGACATGCGCGGAGGAGCTTACTATGGCTCAACCTGCTGATCCCC 720
QY 252 CysValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGlu 271
Db :||| ::||| :::: |||
721 TGTGTGTCATCTCCGCCCTCGCTGTGTTCCTGCTCTGATGCTGCTGCTGCTGCTGCT 780
QY 272 LysLeuSerLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValala 291
Db :||| ::||| :::: |||
781 AAGATTTCCTGGGATAACAGTCTTACTCTCTTACCGTCTTATCGCTCATGCTGCTGCTGCT 840
QY 292 GluThrMetProAlaThrSerAspAlaValProLeuLeuGlyThrTyrPheAsnCysIle 311
Db :||| ::||| :::: |||
841 GAGATCATGCCCGCAACATCCGATTCCGATACCAATGTAGTCCCACTACTTCCGAGCAGC 900
QY 312 MetPheMetValAlaSerSerValValSerThrIleLeuIleLeuAsnThrHisHiArg 331
Db :||| ::||| :::: |||
901 ATGATCATGTGGGCCCTCTCGTGTGTGAGCGGTGATCGTGTGCTGCTGCTGCTGCTGCT 960
QY 332 HisAlaAspThrHisGluMetSerAspTrpIleArgCysValPheLeuTyrTrpLeuPro 351
Db :||| ::||| :::: |||

961 GACCCCGACGGGGCAAGATGCCCAAGTGGACCAAGTGCATCTTCTGAACTGGTGCAG 1020
 QY TrrValLeuArgMetSerArgProGly----- 360
 ||| ||||| |||||
 Db 1021 TGGTTCCTCGGANTGAAGAGCGCCGGGAGGACAAAGTGCGCCCGGCTGCCAGCAACAAG 1080
 ||| ||||| ||||| |||||
 QY 361 -----SerAlaThrThrProProProAlaArg 369
 ||| ||||| ||||| |||||
 Db 1081 CAGCGCGCGCTGCAGCGCTGGCCAGTGTGGAGATGAGCGCCGTGGCGCGCGCCGCCGACG 1140
 ||| ||||| ||||| |||||
 QY 370 -----ValProPro 372
 ||| ||||| ||||| |||||
 Db 1141 AACGGGAACCTGCTGTACATCGGCTTCGCGGCCCTGGACGGCGTGCACTGTGTCCCGACC 1200
 ||| ||||| ||||| |||||
 QY 373 ProProAspLeuGluLeuArgGluArgSerSerLysSerLeuLeuAlaAsnValLeuAsp 392
 ||| ||||| ||||| |||||
 Db 1201 CCCGACTCTGGGTAGTGTGGCGCGCATGGCCGTGCTCCCGCCAGCAC----- 1248
 ||| ||||| ||||| |||||
 QY 393 IleAspAspPheArgHisProGlnAlaGlnProGlnCysCysArgTyrTyrArg 412
 ||| ||||| ||||| |||||
 Db 1249 --GATGAGCACCTCCTGCAC----- 1266
 ||| ||||| ||||| |||||
 QY 413 GlyGlyGluGluAsnGlyAlaGlyLeuAlaAlaHisSerCysPheGlyValAspTyrGlu 432
 ||| ||||| ||||| |||||
 Db 1267 GGTGGCAACCCCCGAGGGG-----GACCCGAC 1296
 ||| ||||| ||||| |||||
 QY 433 LeuSerLeuIleLeuLysGluIleArgValIleThrAspGlnMetArgLysAspAspGlu 452
 ||| ||||| ||||| |||||
 Db 1297 TTGCCCAAGATCCTGGAGGAGTCCGCTACATTGCCAACCGCTTCGCGTCCAGGACGAA 1356
 ||| ||||| ||||| |||||
 QY 453 AspAlaAspIleSerArgAspTrrPylsPheAlaAlaMetValValAspArgLeuCysLeu 472
 ||| ||||| ||||| |||||
 Db 1357 AGCGAGCGGTCTGCAGCGAGTGGAGTTCGCCGCCCTGTGTGTGGACCGCCTGTGCCCTC 1416
 ||| ||||| ||||| |||||
 QY 473 IleIlePheThrLeuPheThrIleIleAlaThrLeuAlaValLeuLeuSerAlaProHis 492
 ||| ||||| ||||| |||||
 Db 1417 ATGGCTTCTCGGTCTTCACCATCATGTGCACCATCGGCATCTGTATGTGGCTCCCAAC 1476
 ||| ||||| ||||| |||||
 QY 493 IleMet 494
 ||| ||||| ||||| |||||
 Db 1477 TTCGTG 1482
 ||| ||||| ||||| |||||
 RESULT 13
 AAC90380
 ID AAC90380 standard; cdna; 1509 BP.
 XX ||||| ||||| |||||
 AC AAC90380;
 XX ||||| ||||| |||||
 DT 14-MAR-2001 (first entry)
 XX ||||| ||||| |||||
 DE Wild-type human alpha7 ligand gated ion channel coding sequence.
 XX ||||| ||||| |||||
 KW Human; alpha7 nicotinic acetylcholine gated ion channel;
 KW 5-hydroxytryptamine; 5-HT3; calcium ion conductance; ss.
 OS Homo sapiens.
 XX ||||| ||||| |||||
 XX W0200073431-A2.
 FN ||||| ||||| |||||
 PD 07-DEC-2000.
 XX ||||| ||||| |||||
 XX 25-MAY-2000; 2000WO-US011862.
 PF ||||| ||||| |||||
 XX 27-MAY-1999; 99US-0136174P.
 PR ||||| ||||| |||||
 XX (PHAA) PHARMACIA & UPJOHN CO.
 XX G-oppi VE, Wolfe ML, Berkenpas MB;
 XX WPI; 2001-061524/07.
 DR P-PSDB; AAB50012.
 XX ||||| ||||| |||||
 PT Special cell culture medium for treating cells and for inducing mammalian

Special cell culture medium for treating cells and for inducing mammalian

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:31:03 ; Search time 12,0702 Seconds

(without alignments)
3952.801 Million cell updates/sec

Title: US-09-303-232-4

Perfect score: 2665

Sequence: 1 MCGRRARRSHLAAPAGLLLL.....LFTTIATLAVLLSAPHIMVS 496

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1267	47.5	502	2 JN0113	nicotinic acetylch
2	1242	46.6	502	2 G02259	alpha 7 neuronal n
3	1241	46.6	502	1 ACHUA7	nicotinic acetylch
4	1241	46.6	502	2 A57175	nicotinic acetylch
5	1232.5	46.2	502	2 T01378	nicotinic receptor
6	1219.5	45.8	511	2 JH0173	alpha-bungarotoxin
7	1201	45.1	498	2 S68588	nicotinic acetylch
8	1141	42.8	461	2 T25671	hypothetical prote
9	1060	39.8	560	2 T19622	hypothetical prote
10	1055.5	39.6	542	2 T19862	hypothetical prote
11	979	36.7	557	2 S12359	nicotinic acetylch
12	962	36.1	503	2 A53956	nicotinic acetylch
13	959	36.0	528	1 ACCH2N	nicotinic acetylch
14	957.5	35.9	567	1 ACFFA1	nicotinic acetylch
15	954	35.8	502	2 A37040	nicotinic acetylch
16	953	35.8	511	2 A40110	nicotinic acetylch
17	936.5	35.1	495	2 S60589	acetylcholine rece
18	932	35.0	499	2 A24572	nicotinic acetylch
19	926	34.7	576	1 ACFFA2	nicotinic acetylch
20	918.5	34.5	456	1 ACCHAN	nicotinic acetylch
21	908.5	34.1	491	1 ACCHNN	nicotinic acetylch
22	904	33.9	622	1 ACCH4N	nicotinic acetylch
23	903.5	33.9	498	2 G02421	nicotinic acetylch
24	902.5	33.9	502	2 S10505	nicotinic acetylch
25	902	33.8	461	2 I50548	nicotinic acetylch
26	901	33.8	503	2 JH0174	acetylcholine rece
27	900.5	33.8	494	2 T03289	nicotinic acetylch
28	899	33.7	461	1 ACRYA1	nicotinic acetylch
29	898.5	33.7	512	2 B37014	nicotinic acetylch

ALIGNMENTS

RESULT 1

JN0113

nicotinic acetylcholine receptor alpha-7 chain precursor, neuronal - chicken
N;Alternate names: alpha-bungarotoxin-binding protein alpha chain
C;Species: Gallus gallus (chicken)

C;Date: 12-Feb-1993 #sequence revision 12-Feb-1993 #text_change 20-Aug-1999
C;Accession: JN0113; JH0172; S28018; B25738; S26566

R;Couturier, S.; Bertrand, D.; Matter, J.M.; Hernandez, M.C.; Bertrand, S.; Millar, N.;

Neuron 5, 847-856, 1990

A;Title: A neuronal nicotinic acetylcholine receptor subunit (alpha 7) is developmental

A;Reference number: JN0113; MUID:91097796; PMID:1702646

A;Accession: JN0113

A;Molecule type: DNA

A;Residues: 1-502 <COU>

A;Cross-references: GB:X68586; NID:G287756; PIDN:CAA48576.1; PID:G287757

A;Experimental source: white leghorn; brain

R;Schoepfer, R.; Conroy, W.G.; Whiting, P.; Gore, M.; Lindstrom, J.

Neuron 5, 35-48, 1990

A;Title: Brain alpha-bungarotoxin binding protein cDNAs and MABs reveal subtypes of thi

A;Reference number: JH0172; MUID:90315158; PMID:2369519

A;Accession: JH0172

A;Molecule type: mRNA

A;Residues: 1-502 <SCH>

A;Cross-references: EMBL:X52295; NID:G63077; PIDN:CAA36543.1; PID:G63078

A;Experimental source: brain

R;Matter-Sadinski, L.; Hernandez, M.C.; Roztocil, T.; Ballivet, M.; Matter, J.M.

EMBO J 11, 4529-4538, 1992

A;Title: Neuronal specificity of the alpha7 nicotinic acetylcholine receptor promoter de

A;Reference number: S28018; MUID:93049204; PMID:1425587

A;Accession: S28018

A;Molecule type: DNA

A;Residues: 1-18 <MAS>

A;Cross-references: EMBL:X68246; GB:S49751; NID:G65319; PIDN:CAA48317.1; PID:G65320

A;Experimental source: white leghorn; erythrocyte

R;Conti-Ironcon, B.M.; Dunn, S.M.J.; Barnard, E.A.; Dolly, J.O.; Lai, F.A.; Ray, N.; Ra

Proc. Natl. Acad. Sci. U.S.A. 82, 5208-5212, 1985

A;Title: Brain and muscle nicotinic acetylcholine receptors are different but homologous

A;Reference number: A94055; MUID:85270494; PMID:3860855

A;Accession: B25738

A;Molecule type: protein

A;Residues: 24-25, 'ET', 28-41, 'X', 43-45, 'X', 47 <CON>

C;Comment: This acetylcholine receptor is blocked by alpha-bungarotoxin and is localized

C;Genetics:

A;Introns: 19/1; 65/3; 80/3; 117/2; 144/1; 200/1; 265/1; 294/1; 330/3

C;Superfamily: acetylcholine receptor

F;Keywords: brain; glycoprotein; ion channel; neurotransmitter receptor; phosphoprotein;

F;24-502/Product: nicotinic acetylcholine receptor predicted <SIG>

F;231-254/Domain: transmembrane #status predicted <TR1>

F;262-280/Domain: transmembrane #status predicted <TR2>

F;296-317/Domain: transmembrane #status predicted <TR3>

F;470-488/Domain: transmembrane #status predicted <TR4>

64 DVDEKNQVLTTNIWLQMSWTDHYLQWNVSEYPGVKTVRFPDGGIWKPKDILLYNSADERFD 123

QY 122 M YNSADEGEFDSTYPTNVVRNNGSCLYVPPGIEKSTCKIDITWFPEDDORCEMKFGSWTY 181

0v 122 M YNSA DEGEDSTYPTNWWRNNGSCLYVPRCTEKSTCKITITWEPEDDOPCEMKECSTWY 101

```

1,112/ training size: phosphate (yr) (covarenc) #status predicted

```


A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-542 <WIL>
A;Cross-references: EMBL:Z70266; PIDN:CAA94206.1; GSPDB:GNO0028; CESP:C40C9.2
A;Experimental source: clone C40C9
C;Genetics:
A;Gene: CESP:C40C9.2
A;Map position: X
A;Insertions: 11/3; 69/3; 123/2; 173/3; 216/1; 248/1; 279/1; 314/1; 350/3; 430/1; 464/1; 49
C;Superfamily: acetylcholine receptor

Query Match 39.6%; Score 1055.5; DB 2; Length 542;
Best Local Similarity 40.5%; Pred. No. 5.6e-84;
Matches 214; Conservative 94; Mismatches 136; Indels 85; Gaps 12;

QY 34 EKRLHLLHLLHYNLVRPVVNESDPLQLSFLGLTLMQIIDVDEKNQLLITNWLKLENDM 93
DB 30 EYRLADLRNLDYERPVANASEFLVSVKYLQQILDVDERNQVITLVAMIEYQWTDY 89

QY 94 NLRWNTSDFGGVKDLRVP--PHRLWKPDVLMYNSADEGFDSTYPTNVVVRNNGSCLYVPP 151
DB 90 KKLWDPSEYGGIKDIRIPGNANALWKPDVLLYNSADENFDSTYPTNVVSVYTGDLQVPP 149

QY 152 GIFKSTCKIDITWFPDDQRCCKMFGSWTGYQLDLQLQ-----DEGGDISSVFTN 204
DB 150 GILKLSCKIDITYPPDDQICHLKFGSWTYSGNFIDLRINGPEGKNISDEGIDVQYVQN 209

QY 205 GEWELIGVPGKREIYVNCPPERYDITFAVIRRKTLVYFFENLIVPCVLIASMALLGFT 264
DB 210 GEWNLAVPARHETNFD--EQYPSLFFLIQIRRTLYGLNLIIFSLISLMTVIGFT 267

QY 265 LPDPSGKLSGLVTILSLVFLMVAETMPATSDAVPELL-----GTYFNCIMFMVASS 318
DB 268 LPDAGEKITLEITILSVCFLLSMVADMTPTSEAVPLIGLIIFSGAFTSCCMLVVSAS 327

QY 319 VVSTILILNHYHRRHADTHEMSDWIRCVFLYWL PWLVRMSRPG-----SATTPPPA 368
DB 328 WFTVLVNLNHNKRPETHMSPEFLRELLIWL PWLLLMRRPGKTFNCTHLKAEKAEKA 387

QY 369 RV-----PPPDLELRSSKSLANV-----LDIDDDF--RH-----PQAOQ 404
DB 388 KQSIKNGVCGKFTSDVHPSGSLMKNTKLGRQOIIDFEYFHVQHNLMFVAPSEMT 447

QY 405 PQCC-----RYRGGENGAGLAHSCFGVDYELSLILKEIRVITD 445
DB 448 PRVTYSKMAESYVEDVMTLNKYMQ-----KACL-----ELKNISSQTRAMRK 492

QY 446 QMRKDEDEDADISRDWKFAMVVDRLCLIIITLTIATLAVLISAPHIM 494
DB 493 KMEDEDERQAAANDWKFAMVVDRLCLITTFVSTGIMFSSPHLI 541

RESULT 11
S12359
nicotinic acetylcholine receptor alpha-L1 chain precursor - desert locust
C;Species: Schistocerca gregaria (desert locust)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 20-Aug-1999
C;Accession: S12359
R;Marshall, J.; Buckingham, S.D.; Shingai, R.; Lunt, G.G.; Goosey, M.W.; Darlison, M.G.;
EMBO J. 9, 4391-4398, 1990
A;Title: Sequence and functional expression of a single alpha subunit of an insect nicot
A;Reference number: S12359; MUID:91092263; PMID:1702381
A;Accession: S12359
A;Molecule type: mRNA
A;Residues: 1-557 <MAR>
A;Cross-references: EMBL:X55439; NID:g10133; PIDN:CAA39081.1; PID:g10134
C;Superfamily: acetylcholine receptor
C;Keywords: Glycoprotein; Ion channel; neurotransmitter receptor; transmembrane protein
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-557/Product: nicotinic acetylcholine receptor alpha-L1 chain #status predicted <MAT
F;245-266/Domain: transmembrane #status predicted <TM1>
F;274-295/Domain: transmembrane #status predicted <TM2>
F;308-329/Domain: transmembrane #status predicted <TM3>

F;501-523/Domain: transmembrane #status predicted <TM4>
F;47,235/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 36.7%; Score 979; DB 2; Length 557;
Best Local Similarity 40.0%; Pred. No. 2.8e-77;
Matches 217; Conservative 88; Mismatches 160; Indels 78; Gaps 10;

QY 11 AAPAGLILLCLLWPGRCARCYHEKRLHLLHLLHYNLVRPVVNESDPLQLSFLGLTLMQI 70
DB 4 ALPFPMLILLLLHHPAAANPDAKRLYDILLSNYNRLIRPVSNNTTVLVKGLRSLQ 63

QY 71 IDVDEKNQLLITNWLKLENDMNLNWTSDFGGVKDLRVPVPPHPLKVPDVLMYNSADEGF 130
DB 64 IDLNLKQILTTNWLSEHQDKHFRWDPAEYGGVTELYVSEHILWLPDILVLYNADGEY 123

QY 131 DSYPTNVVVRNNGSCLYVPPGIFKSTCKIDITWFPDDQRCCKMFGSWTGYQLDL-- 188
DB 124 VVTMTKAVLHHTGKVVMTTPAIFKSSCEIDVRYFPDQQTFCFMKFGSWTGYDQIDLKH 183

QY 189 --QLQDEG----GGDISSVFTNGEWELIGVPGKREIYVNCPPERYDITFAVIRRKTL 242
DB 184 INQYDDNKVKVGLDLREYVPSVENDILGVPAERHEKYPCCAEPYDIFENITLRKTL 243

QY 243 YFFNLIVPCVLIASMALLGFTLPPDSGKLSGLVTILSLTVFLNMVAETMPATSDAVP 302
DB 244 FYTNLIVPCVIGISYLSVLVFLPADSGEKIALCISILLQTMFFLLISEIIPSTSLAP 303

QY 303 LLGYTFNCIMFMVASSVSVTILINHYHRRHADTHEMSDWIRCVFLYWL PWLVRMSRPGSA 362
DB 304 LLGKYLFTVVLVGLSVITIMVLNHYRKPSTHKMAPVWRKVIRLPLKILLM----- 357

QY 363 TTPPPARVPPPPDLRLRERSKSL-----ANVLIDDDDFRHPQ 401
DB 358 -----RV---PEQLADLASKRLLRHANSKLSAAAAAASASSPSDLRHH 408

QY 402 AQDPQCCRY-----YRGGENGAGLAH-----SCFGVD----- 430
DB 409 LHQHQHQLHLQHLQPGCGNLHSATNFGGSAGAFGLPSVVGDLGSLSDVATRKKY 468

QY 431 -YELSLILKEIRVITDOMRKDE-DADISRDWKFAMVVDRLCLIIITLTIATLAVL 488
DB 469 PFELEKAHNVLFIQNHMQRQDFDAB-DQDWGFVAMVLDRLFLWIFTIASIVGTFAILC 527

QY 489 SAP 491
DB 528 EAP 530

RESULT 12
A53956
nicotinic acetylcholine receptor alpha-3 chain precursor, neuronal - human
C;Species: Homo sapiens (man)
C;Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 20-Aug-1999
C;Accession: A53956; S21338
R;Mihovilovic, M.; Roses, A.D.
Exp. Neurol. 111, 175-180, 1991
A;Title: Expression of mRNAs in human thymus coding for the alpha3 subunit of a neuronal
A;Reference number: A53956; MUID:91114756; PMID:1989896
A;Accession: A53956
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-503 <MIH>
A;Cross-references: GB:M37981; NID:g189252; PIDN:AAA59942.1; PID:g189253
R;Anand, R.; Lindstrom, J.
submitted to the EMBL data Library, June 1990
A;Description: Nucleotide sequence of the mature human nicotinic acetylcholine receptor
A;Reference number: S21338
A;Accession: S21338
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 30-503 <ANA>
A;Cross-references: EMBL:X53559; NID:g34985; PIDN:CAA37625.1; PID:g34986
C;Genetics:

A:Gene: GDB:CHRNA3
A:Cross-references: GDB:125219; OMIM:118503
A:Map position: 15q24-15q24
C:Superfamily: acetylcholine receptor
C:Keywords: neurotransmitter receptor

Query Match 36.1%; Score 962; DB 2; Length 503;
Best Local Similarity 41.1%; Pred. No. 7.3e-76;
Matches 206; Conservative 89; Mismatches 172; Indels 34; Gaps 11;
QY 10 LAAPAG---LILLICLLPRGARGCYHEKRLHLHLLDHYNVLRLPVPVNESDPLQLSFG 65
DB 6 LAAPGAVARLILLLSLIP- VARASEAERLPERLFEDYNEIIRPVANVSDPVIIHFV 64
QY 66 TLQIIVDEKKNOLLITNWLKLEWDMNMLRWNTSDFGVKDLRVPPHRLMKDPDVLMYNS 125
DB 65 SMSQLVKVDEVNQIMETNLWLKQIMNDYKLVKWPSPDYGGAERKRVPAQKIWKPDIVLYNN 124
QY 126 ADEGFDSTYPTNVVVRNNGSLYVPPGIKSTCKIDITWFPDDORCEMKFGSWTYDGYQ 185
DB 125 AVGDFOVDKTKALLKYTGVTWIPPAIFKSSCKIDITVFPFDYQNTWKFGSWYDKAK 184
QY 186 LDQLQDEGGDISSTFWNGEWELIGVPGKRNEIYVNCPCPEYIDITFAVIRRKTYLF 245
DB 185 IDLVIG-SSMNLKDYWSEGEWAILKAPGYKHIDKNCCEIYPDITYSLYIRRLPLFYT 243
QY 246 FNLIVPCVLIIASNALGFTLPDPSGKLSGVITLLSLTVFLNMVAETMPATSDAVPLLG 305
DB 244 INLIIPCLLSIFLTVLVFVLPSCGKVKVLCISVLLSLTVFLVITETIPSTSLVPIIG 303
QY 306 TYFNCITFMVAVSSVSTILINVHRHADPHENSDWIRCVLYLWLPWLMRSP- - - - -GS 361
DB 304 EYLLTFMIFVTLISIVTVFLNVHYRTPTHTTWSWKTVFLNLLPRVFMPTREPTNEGN 363
QY 362 ATTPPARVPPPPPLELRLRS- SKSLANLVLDIDDDFRHPQAQPOC- - - - -CRYRGGEEN- 417
DB 364 AQPRFLYGAELNLCNCFSAESKGKEG- - - - -YP-CDQMGYCHHRIKISNF 413
QY 418 GAGLA-AHSGFGVDYELSL- - - - -ILKEIRVITDMQRKDDADISRDKFAAMVV 467
DB 414 SANLTRSSSESVDVLSLSALSPEIKAEIQSVKYIAENKMAQNEAKEIQDDWKYVAMVI 473
QY 468 DRCLLIIFTLFTIATLAVLL 488
DB 474 DRIFLWVFTLVCLGTAGLFL 494

RESULT 13

ACCH2N

nicotinic acetylcholine receptor alpha-2 chain precursor, neuronal - chicken
C:Species: Gallus gallus (chicken)
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 21-Jul-2000
C:Accession: S00377
R:Nef, P.; Oneyser, C.; Alliod, C.; Couturier, S.; Ballivet, M.
EMBO J. 7, 595-601, 1988
A:Title: Genes expressed in the brain define three distinct neuronal nicotinic acetylcholine receptor subunits
A:Reference number: S00376; MUID:88283624; PMID:3267226
A:Accession: S00377
A:Molecule type: DNA
A:Residues: 1-528 <NEFF>
A:Cross-references: EMBL:X07339; NID:g62792; PIDN:CAB59645.1; PID:g6136914
C:Genetics:
A:Introns: 21/1; 73/3; 88/3; 125/2; 487/3
C:Superfamily: acetylcholine receptor
C:Keywords: glycoprotein; ion channel; neurotransmitter receptor; postsynaptic membrane;
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-528/Product: nicotinic acetylcholine receptor alpha-2 chain #status predicted <MAT>
F:241-264/Domain: transmembrane #status predicted <TM1>
F:272-290/Domain: transmembrane #status predicted <TM2>
F:306-327/Domain: transmembrane #status predicted <TM3>
F:502-520/Domain: transmembrane #status predicted <TM4>
F:54, 104/Binding site: carboxyrate (Asn) (covalent) #status predicted
F:158-172, 222-223/Disulfide bonds: #status predicted

Query Match 36.0%; Score 959; DB 1; Length 528;
Best Local Similarity 39.2%; Pred. No. 1.4e-75;
Matches 198; Conservative 84; Mismatches 161; Indels 62; Gaps 10;

QY 31 GYHEKRLHLHLLDHYNVLRLPVPVNESDPLQLSFGTLQIIVDEKKNOLLITNWLKLEW 90
DB 31 GFAERLFLKHLFTGYNRWSRPVNTSDVIVVKEGLSLAQIIVDEKKNOMTINWLKQEW 90
QY 91 NDMNLRWNTSDFGGVKDLRVPPHRLMKDPVLMYNSADEGFDSTYPTNVVVRNNGSLYVP 150
DB 91 SDYKLRWNPEDFDNVTISIRVSEMIWIPDILVYNNADGEFAVTHMTKAHLFSNGKVKWVP 150
QY 151 PGIFKSTCKIDITWFPDDORCEMKFGSWTYDGYQLDQLQDEGGDISSTFWNGEWELI 210
DB 151 PAIYKSSCSIDITYFPFDQCNCKMFGSWTYDKAKIDLE-NMEHHVDLKDYESEGEWAIL 209
QY 211 GYVGRKNEIYVNCPCPEYIDITFAVIRRKTYLFVFNLIYVPCVLIIASMALGFTLPDPSG 270
DB 210 NAIGRYSKKYDCCYEIYFDITFYVIRRLPLFYTNLIIPCLLSISLTVLVFVLPDSCG 269
QY 271 EKLSLGVITLLSLTVFLNMVAETMPATSDAVPLLGTYFNCIMPMVAVSSVSTILINVHH 330
DB 270 EKITLCISVLLSLTVFLVITETIPSTSLVPIIGSYLLFTMIFVTLISIIITVFLNVHH 329
QY 331 RHADTHENSDWIRCVLYLWLPWLMRSPGSAATPPPARVP- - - - -PPDLELRERS 382
DB 330 RSPSTHMPHWRSFPLGFIPRLFMKR- - - - -PPLLLPAEGTTGQVDPGTRL- - - - -ST 380
QY 383 SKSLANLVLDIDDDFRHPQAQPOC- - - - -CCRYRGGEENGAGLAH-SC- - - - -426
DB 381 SRCWLET- - - - -DVDDKWEEREEEEEEEEEEEEKAYSRVPSGSGQ- - - - -GTQCHYSERQA 435
QY 427 - - - - -FGVDYELSL- - - - -ILKEIRVITDMQRKDDADISRDKFA 463
DB 436 GKASGGPAQVPLKGEVSGDQGLTSLPSILRALEGVQVIADHLRAEDADFSVKEDWKYV 495
QY 464 AMVVDRLCLLIIFTLFTIATLAVLL 488
DB 496 AMVIDRIFLWMTIIVCLLGTAGLFL 520

RESULT 14

ACFPAL

nicotinic acetylcholine receptor alpha-1 chain precursor - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 21-Jun-2002
C:Accession: S00381; A38801
R:Bosny, B.; Ballivet, M.; Spierer, P.
EMBO J. 7, 611-618, 1988
A:Title: Conservation of neural nicotinic acetylcholine receptors from Drosophila to vertebrates
A:Reference number: S00381; MUID:88283626; PMID:2840281
A:Accession: S00381
A:Molecule type: DNA
A:Residues: 1-567 <BOS>
A:Cross-references: GB:X07194; NID:g7575; PIDN:CAA30172.1; PID:g7576
A:Accession: A38801
A:Molecule type: mRNA
A:Residues: 1-567 <BOS>
A:Cross-references: EMBL:X07194; NID:g7575; PIDN:CAA30172.1; PID:g7576
A:Note: 538-Tyr was also found
C:Genetics:
A:Gene: FlyBase:ncr-alpha-96Aa
A:Cross-references: FlyBase:FBgn0000036
A:Map position: 3R 96A
A:Introns: 64/3; 79/3; 116/2; 176/3; 330/2; 401/1; 499/3
C:Superfamily: acetylcholine receptor
C:Keywords: glycoprotein; ion channel; neurotransmitter receptor; postsynaptic membrane
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-567/Product: nicotinic acetylcholine receptor alpha-like chain #status predicted <F>
F:22-240/Domain: extracellular #status predicted <EXT>
F:240-264/Domain: transmembrane #status predicted <TM1>
F:272-290/Domain: transmembrane #status predicted <TM2>

QY
Db

10 LAAPAGLLILLCLLWPRGARGCYHEKRLHHLLDHYNVLERPVVNESDPLQLSFLTLMQ 69
||| ||| ||| ||| ||| : ||| : ||| : ||| : ||| :
9 LACRARLLLLLSLP- VARASAEHRLFELFEDYNEIIRPVANVSDEVIIEFVSMQ 67

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:18:28 ; Search time 9.26316 Seconds

(without alignments)
2788.120 Million cell updates/sec

Title: US-09-303-232-4

Perfect score: 2665

Sequence: 1 MGRARRSHLAAPAGALLLL.....LFTIITLAVLLSAPHIMVS 496

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1267	47.5	502	1	ACH7_CHICK
2	1246	46.8	502	1	ACH7_HUMAN
3	1241	46.6	502	1	ACH7_MOUSE
4	1232.5	46.2	502	1	ACH7_RAT
5	1218.5	45.7	499	1	ACH7_BOVIN
6	1201	45.1	498	1	ACH1_CABEL
7	980.5	36.8	516	1	ACH1_MANSE
8	979	36.7	557	1	ACH1_SCHGR
9	960	36.0	503	1	ACH3_HUMAN
10	960	36.0	529	1	ACH2_HUMAN
11	959	36.0	528	1	ACH2_CHICK
12	956.5	35.9	567	1	ACH1_DROME
13	951	35.7	511	1	ACH2_RAT
14	936.5	35.1	495	1	ACH3_BOVIN
15	930	34.9	499	1	ACH3_RAT
16	926	34.7	576	1	ACH2_DROME
17	918.5	34.5	456	1	ACHA_CHICK
18	918	34.4	434	1	ACH6_HUMAN
19	917	34.4	496	1	ACH3_CHICK
20	908.5	34.1	491	1	ACHN_CHICK
21	906	34.0	461	1	ACHA_TORMA
22	904	33.9	622	1	ACHA_CHICK
23	903.5	33.9	498	1	ACHP_HUMAN
24	902.5	33.9	502	1	ACHN_HUMAN
25	902	33.8	495	1	ACHP_RAT
26	899	33.7	461	1	ACHA_TORCA
27	898.5	33.7	512	1	ACH3_CARAU
28	898	33.7	500	1	ACHN_RAT
29	896.5	33.6	457	1	ACHA_BOVIN
30	896.5	33.6	493	1	ACH6_RAT
31	893.5	33.5	519	1	ACH4_DROME
32	890	33.4	457	1	ACHA_RAT
33	889	33.4	457	1	ACHA_MOUSE

34	889	33.3	470	1	ACHP_CHICK
35	886.5	33.3	494	1	ACH6_CHICK
36	885.5	33.2	459	1	ACHN_CARAU
37	880	33.0	456	1	ACHA_BRARE
38	879.5	33.0	627	1	ACH4_HUMAN
39	876.5	32.9	457	1	ACH2_XENLA
40	874	32.8	521	1	ACH3_DROME
41	872	32.7	538	1	ACH8_CABEL
42	867.5	32.6	630	1	ACH4_RAT
43	866.5	32.5	482	1	ACHA_HUMAN
44	865.5	32.5	629	1	ACH4_MOUSE
45	858.5	32.2	452	1	ACH5_RAT

ALIGNMENTS

RESULT 1
ACH7_CHICK
ID ACH7_CHICK STANDARD; PRT; 502 AA.
AC P22770;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neuronal acetylcholine receptor protein, alpha-7 chain precursor.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=91097796; PubMed=1702646;
RA Couturier S., Bertrand D., Matter J.M., Hernandez M.C., Bertrand S.,
RA Millar N., Valera S., Barkas T., Ballivet M.;
RT "A neuronal nicotinic acetylcholine receptor subunit (alpha 7) is
RT developmentally regulated and forms a homo-oligomeric channel blocked
RT by alpha-BTX.";
RL Neuron 5:847-856(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=90315158; PubMed=2369519;
RA Schoeffer R., Conroy W.G., Whiting P., Gore M., Lindstrom J.;
RT "Brain alpha-bungarotoxin binding protein cDNAs and MABs reveal
RT subtypes of this branch of the ligand-gated ion channel gene
RT superfamily";
RL Neuron 5:35-48(1990).
RN [3]
RP SEQUENCE OF 1-18 FROM N.A.
RC STRAIN=White leghorn; TISSUE=Erythrocyte;
RX MEDLINE=93049204; PubMed=1425587;
RA Matter-Sadzinski L., Hernandez M.C., Roztocil T., Ballivet M.,
RA Matter J.M.;
RT "Neuronal specificity of the alpha 7 nicotinic acetylcholine receptor
RT promoter develops during morphogenesis of the central nervous
RT system.";
RL EMBO J. 11:4529-4538(1992).
RN [4]
RP SEQUENCE OF 24-47.
RC TISSUE=Brain;
RX MEDLINE=85270494; PubMed=3860855;
RA Conti-Tronconi B.M., Dunn S.M.J., Barnard E.A., Dolly J.O., Lai F.A.,
RA Ray N., Rafferty M.A.;
RT "Brain and muscle nicotinic acetylcholine receptors are different but
RT homologous proteins.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:5208-5212(1985).
RN [5]
RP MUTAGENESIS OF LEU-270.
RX MEDLINE=92049732; PubMed=1719423;
RA Recah F., Bertrand D., Galzi J.-L., Devillers-Thiery A., Mulle C.,
RA Hussy N., Bertrand S., Ballivet M., Changeux J.-P.;

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; X393604; CAA53802.1; --
DR InterPro; IPR006029; Neur_chan_memb.
DR InterPro; IPR006202; Neur_chan_LBD.
DR InterPro; IPR006201; Neur_chan.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_memb; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMS; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
KW Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
KW Transmembrane; Multigene family; Alternative splicing.
FT SIGNAL 1 19
FT CHAIN 20 499
FT DOMAIN 20 227
FT TRANSMEM 228 252
FT TRANSMEM 259 277
FT TRANSMEM 293 314
FT DOMAIN 315 466
FT TRANSMEM 467 487
FT DISULFID 147 161
FT DISULFID 209 210
FT CARBOHYD 43 43
FT CARBOHYD 87 87
FT CARBOHYD 130 130
FT VARSPLIC 262 290
FT SEQUENCE 499 AA; 56002 MW; AEE5D0B3820D42D5 CRC64;
Query Match 45.78; Score 1218.5; DB 1; Length 499;
Best Local Similarity 46.18; Pred. No. 1e-94;
Matches 241; Conservative 80; Mismatches 115; Indels 87; Gaps 9;
QY 26 RGARC-----GYHEKRLHLHLLDHYNLERPVVNSDDPLQLSGLTLMQIID 72
DB 2 RGSICLALASILHVSILQGEFQKLYKDLVKNYFLERPANDSLPTVVFSLSLQIMD 61
QY 73 VDEKQQLITNIMKLEWDMNLRWNTSDFGGVKDLRVPHRLWKPDVLMVNSADEGDS 132
DB 62 VDEKQVLTNIMLQMTWTHTYLOWNASPYGVKTVRFPDQGIWKPDIILYNSADERFDA 121
QY 133 TYPTNVVRNGSLVYPPGIFKSTCKIDITWRPFDDQRCMKFGSWTYDGYQLDLQLOD 192
DB 122 THTNVLNVSSGHGCOYLPGLFKSCVIDWRPFDDVQCKLKFGSWYGGWSLDLQOE 181
QY 193 EGGGDISFVTNGEWELIGVPGKENEIYNCCPEYIDITFVAVIRKTLTYFFNLVIPC 252
DB 182 ---ADISGYIPNGEMDLVGLGKSEKFEYCKEPPDVTFTVSIRRTLYYGLNLLIPC 238
QY 253 VLTASMLLGTLPDPSGKSLGVTLLSLTVPFLNVAETMPATSDAVPLLGYFNCIM 312
DB 239 VLISALALLVFLPADSGEKISLGTITVLLSLTVPFLNVAETMPATSDVPLIAYFASTM 298
QY 313 FMVASSVSTLIILNYHHRHADTHMSDTRCFLYWLPLVLRMRSPQ----- 360
DB 299 IIVGLSVVTVIVLYQHHPDGGKMPKATRVULLNWCAMFLMKRGEDKVRPACQHNE 358
QY 361 -----SATTPPPA-----RVPPPPDLRLRSSKSLLANVL--- 391
DB 359 RRCSLASVMSAVGPPATPNGLLYIGRGLDTHMCAPTDP-----SGVVCGRVACSP 411
QY 392 DIDDDFRPQAOQCCEYRGGGEENGAGLAHSCFGVDYVLSLILKEIRVITQMRKDD 451

DB 412 TDEHLLH--AQGP-----SEG-----DPLAKILLEVRYIAHRPCQD 448
QY 452 EDADISRDWKFAMVVDRLCLIFTFTIATLAVLLSAPHIM 494
DB 449 ESEAVCSEWKFACVVDRLCLMAFVFTILCTIGILMSAPNFV 491

RESULT 6
ACH1 CAEEL
ID ACH1 CAEEL STANDARD; PRT; 498 AA.
AC P48180;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Acetylcholine receptor like protein, alpha-type chain precursor.
GN R25G6.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=96196478; PubMed=8627624;
RA Ballivet M., Alliod C., Bertrand S., Bertrand D.;
RT "Nicotinic acetylcholine receptors in the nematode Caenorhabditis
RT elegans.";
RL J. Mol. Biol. 258:261-269(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=96196478; PubMed=8627624;
RA Nelson J., Wohldmann P.;
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Possible acetylcholine receptor.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: Belongs to the ligand-gated ionic channel family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

EMBL; X83887; CAA58764.1; --
EMBL; AF022973; AAC25796.1; --
PIR; S68588; S68588.
HSP; P58154; I19B.
DR WormPep; F25G6.3; CE09639.
DR InterPro; IPR006029; Neur_chan_memb.
DR InterPro; IPR006202; Neur_chan_LBD.
DR InterPro; IPR006201; Neur_chan.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_memb; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMS; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
KW Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein;
KW Transmembrane; Signal.
FT SIGNAL 1 19
FT CHAIN 20 498
FT DOMAIN 20 230
FT TRANSMEM 231 252
FT TRANSMEM 261 279
FT TRANSMEM 295 314
FT DOMAIN 315 472
FT TRANSMEM 473 493
FT DISULFID 147 161
FT DISULFID 211 212
FT CARBOHYD 43 43
FT POTENTIAL.
FT ACETYLCHOLINE RECEPTOR LIKE PROTEIN,
FT ALPHA-TYPE CHAIN.
FT EXTRACELLULAR (POTENTIAL).
FT POTENTIAL.
FT POTENTIAL.
FT POTENTIAL.
FT CYTOPLASMIC (POTENTIAL).
FT POTENTIAL.
FT BY SIMILARITY.
FT ASSOCIATED WITH RECEPTOR ACTIVATION
FT (BY SIMILARITY).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT Missing (in isoform Short).
FT /FTid=VSP_000075.
FT SEQUENCE 499 AA; 56002 MW; AEE5D0B3820D42D5 CRC64;


```
Db 149 NNADGEFAVTHMTKAHLFSTGTGVHVPVPAIYKSCSIDVTFPPDQCNCKWKGSTYDK 208
QY 184 YQLDLQLOBGGDIDSSFYVNGEWELLGVGKNEIYVNCPPYDITFAVVRKTL 243
Db 209 AKIDLE-QMEQTVLDKYWESGWAIVNATGTNYSKYDCCAIYDPVYAFVIRLPLF 267
QY 244 YFFNLVPCVLIASMAALLGFTLPDSCGKLSGLVTLISLTVFLNMVAETMPATSDAVPL 303
Db 268 YTINLIIPCLLISCLTVLYFLVSDCGEKTLCISVLLSTVFLLLITLIPSTSLVPL 327
QY 304 LGFYNCIMFVASSVSTILINYYHHRHADTHEMSDWIRCVLYMLPWVLRMSRGSAT 363
Db 328 IGEYLLFTMTFVTLISVITVFLNVHRSPTSHTMPHWVRGALLGCVPRLLMNRP--- 383
QY 364 TPPARVPPPPDLELRSSKSLIANVLD-----IDDDF----- 398
Db 384 -PPFVELCHP--URLKLSFSYHLENSVDAEREVVVEEDRWACAGHVAPSVTILCSHG 440
QY 399 --HPOAQQPCQCRYRGEENGAGLAHSCFGVDYELSLILKEIRVITDQMRKDDADADI 456
Db 441 HLHSGASGPKAEALLQEGE---LLSPH-----MQKALEGVHYIADHLRSEDADSSV 489
QY 457 SRDWKFAAMVVBRLCLIIITFTIATLAVLL 488
Db 490 KEDWKYVAMVIDRIFLWLFIVCFGLGTIGLFL 521

RESULT 11
ACH2_CHICK
ID_ACH2_CHICK STANDARD; PRT; 528 AA.
AC P09480;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Neuronal acetylcholine receptor protein, alpha-2 chain precursor.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]_TaxID=9031;
SEQUENCE FROM N.A.
RP TISSUE=Brain;
RC MEDLINE=88283624; PubMed=3267226;
RA Nef P., Onysier C., Alliod C., Coururier S., Ballivet M.;
RT "Genes expressed in the brain define three distinct neuronal
RT nicotinic acetylcholine receptors.";
RL EMBO J. 7:595-601(1988).
CC -!- FUNCTION: After binding acetylcholine, the AChR responds by an
CC extensive change in conformation that affects all subunits and
CC leads to opening of an ion-conducting channel across the plasma
CC membrane.
CC -!- SUBUNIT: Neuronal AChR seems to be composed of two different
CC type of subunits: alpha and non-alpha (also called beta). A
CC functional receptor seems to consist of two alpha-chains and
CC three non-alpha chains.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X07339; CAB59645.1; -.
CC DR EMBL; X07340; CAB59645.1; JOINED.
CC DR EMBL; X07341; CAB59645.1; JOINED.
CC DR EMBL; X07342; CAB59645.1; JOINED.
CC EMBL; X07343; CAB59645.1; JOINED.
CC DR EMBL; X07344; CAB59645.1; JOINED.
```

```
DR EMBL; AJ250360; CAB59625.1; -.
DR PIR; S00377; ACH2N.
DR InterPro; IPR006029; Neu_channel memb.
DR InterPro; IPR006202; Neur_chan IBD.
DR InterPro; IPR006201; Neur_channel.
DR Pfam; PF02931; Neur_chan IBD; 1.
DR Pfam; PR02932; Neur_chan memb; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMS; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
KW Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
KW Transmembrane; Multigene family.
FT SIGNAL 1 23
FT CHAIN 24 528
FT DOMAIN 24 239
FT TRANSMEM 240 264
FT TRANSMEM 272 290
FT TRANSMEM 306 327
FT DOMAIN 328 501
FT TRANSMEM 502 520
FT DOMAIN 394 410
FT DISULFID 158 172
FT DISULFID 222 223
FT CARBOHYD 54 54
FT CARBOHYD 104 104
FT SEQUENCE 528 AA; 60675 MW; E76C6360AF876364 CRC64;
SQ
Query Match 36.0%; Score 959; DB 1; Length 528;
Best Local Similarity 39.2%; Pred. No. 7.2e-73;
Matches 198; Conservative 84; Mismatches 161; Indels 62; Gaps 10;
QY 31 GVHEKPLHLHLHDHYNVLERPVVNESDPLQLSFGLTLMQIIDVDEKNQLITINIKLEW 90
Db 31 GFAEDRLFKHLFTGYNRWSRPVNTSDVIVVFGLSIAQLIDVDEKNQMTINVLKQEW 90
QY 91 NDMNLWNTSDGGVYKDLVPPHRLKPDVLMVNSADEGFDSTYPTNVVVRNNGSLYVP 150
Db 91 SDYKLRNPEDFDNVTISRVPSEMIIPDIVLYNNADGEFAVTHMTKAHLFSGNKVKKVP 150
QY 151 PGIFKSTCKIDITWPPDDQRCMKFGSWTYDGYQLDLOLBGGGDISSFVTNGEWELI 210
Db 151 PAIYKSCSIDVTFPPDQCNCKWKGSTYDKAKIDLE-NMEHHVDLKDYESGEWALI 209
QY 211 GVPGRKNEIYVNCPPYDITFAVVRKTLIYFFNLVPCVLIASMAALLGFTLPDSCG 270
Db 210 NAIGRYSKKYDCCTEYDITFYFVIRLPLFYTNLIIPCLLISCLTVLYFLVPSDCG 269
QY 271 EKLGLVTLISLTVFLNMVAETMPATSDAVLLGTYFNCIMFVASSVSTILINYYH 330
Db 270 EKITLCISVLLSLTVFLLLITLIPSTSLVILPSTSLVILPSTSLVILPSTSLVIL 329
QY 331 READTHEMSDWIRCVLYMLPWVLRMSRGSATTPPPARVP-----PPDLELRERS 382
Db 330 RSPSTHTPHWVRSFPLGFIPLFMKR-----PPLLLPAEGTQYDPPGTRL--ST 380
QY 383 SKSLIANVLDIDDDFRHPQAQQPQ-----CCRYRGEENGAGLAH-SC----- 426
Db 381 SRCWLET--DVDDKWEESEEEEEEEEEKAYPSKRVPSGSGQ---GTQCHYSCERQA 435
QY 427 -----FGVDYELSI-----ILKEIRVITDQMRKDDADADI SRDWKFA 463
Db 436 GKASGGAPQVPLKGEVSGSQGLTSPSLRALEGVQVIADHLRAEDADFVSKEDWKYV 495
QY 464 AMVDRCLCLIIFTLTIIATLAVLL 488
Db 496 AMVIDRIFLWVFIIVCLLTGVLGL 520
RESULT 12
ACH1_DROME
ID_ACH1_DROME STANDARD; PRT; 567 AA.
```

AC P09478; Q9VC74;
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Acetylcholine receptor protein, alpha-like chain 1 precursor.
 GN NACR-ALPHA-96AA OR ACRB OR ALS OR ACR96AA OR CG5610.
 OS Drosophila melanogaster (Fruit fly).
 OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Oregon-R;
 RX MEDLINE=88283626; PubMed=2840281;
 RA Bossy B., Ballivet M., Spierer P.;
 RT "Conservation of neural nicotinic acetylcholine receptors from
 RT Drosophila to vertebrate central nervous systems.";
 RL EMBO J. 7:611-618(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkely;
 RX MEDLINE=20196006; PubMed=107311132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Franckoch C., Baldwin D.,
 RA Ballou R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Borchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Folsler G., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.B., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.C., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 RN [3]
 RP REVISIONS.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Battencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 RT systematic review.";

Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
 CC -!- FUNCTION: After binding acetylcholine, the AChR responds by an
 CC extensive change in conformation that affects all subunits and
 CC leads to opening of an ion-conducting channel across the plasma
 CC membrane.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: CNS in embryos.
 CC -!- DEVELOPMENTAL STAGE: Late embryonic, late pupal and second instar
 CC larvae stages.
 CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X07194; CAA30172.1; -;
 DR EMBL; AE003747; AAF56301.2; -;
 DR PIR; S00381; ACFEAL;
 DR FlyBase; FBgn0000036; nAChR-alpha-96Aa.
 DR InterPro; IPR006029; Neu_chan_memb.
 DR InterPro; IPR006202; Neur_chan_LBD.
 DR InterPro; IPR006201; Neur_chan.
 DR Pfam; PF02931; Neur_chan_LBD; 1.
 DR Pfam; PF02932; Neur_chan_memb; 1.
 DR PRINTS; PR00252; NRIONCHANNEL.
 DR TIGRFAMS; TIGR00860; LIC; 1.
 DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
 DR Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
 KW Transmembrane; Multigene family; Polymorphism.
 KW SIGNAL 1 21
 FT CHAIN 22 567 ACETYLCHOLINE RECEPTOR PROTEIN, ALPHA-
 FT LIKE CHAIN 1.
 FT DOMAIN 22 240 EXTRACELLULAR.
 FT TRANSMEM 241 264
 FT TRANSMEM 272 290
 FT TRANSMEM 306 325
 FT DOMAIN 326 513 CYTOPLASMIC.
 FT TRANSMEM 514 532
 FT DISULFID 149 163
 FT DISULFID 222 223
 FT CARBOHYD 45 45
 FT CARBOHYD 233 233
 FT VARIANT 538 538
 FT CONFLICT 108 108 L -> H (IN REF. 1).
 SQ SEQUENCE 567 AA; 64019 MW; 08E1F721FB2A92AC CRC64;
 Query Match 35.9%; Score 956.5; DB 1; Length 567;
 Best Local Similarity 38.2%; Pred. No. 1.3e-72;
 Matches 206; Conservative 86; Mismatches 174; Indels 73; Gaps 11;
 Qy 18 LLECLLPARGAGCYHEKRLHLLHLLHYNLRPVVNESDPLQLSFGLTLMQIDVDEKN 77
 Db 9 VFTALHPATGGLANPDAAKRLDYDLLSNLYNLRIPVGNNSDRLTVKMGRLSLQDLVNLKN 68
 Qy 78 QLLITNLTWLEWMDNMLRWNTSDFGCVKDLRVPVPHLWKPDLVLMYNSADEGFDSTYPTN 137
 Db 69 QIMTNVWVEQWWDYKLNWPDYGGVDTLHPVSEHILWPLDIVLYNNAAGNVEVITMTK 128
 Qy 138 VVVRNNGSLYVPPGIFPKSTCKIDITWFPDPDQRCMKFGSWTYDGYQLDL---QLQD- 192
 Db 129 AILHHTGKVMKPPAIFYKSCFIDEVYFDEQTCFMKFGSWTYDGYWDLRLHKLQATDS 188
 Qy 193 ---EGGDISFVTNGEWELIGVPGKNEYNYNCCPEPYDITFPAVIRKILYFFNLI 249
 Db 189 DNIEVGIDLDQYIISVEWDMIRVPAVNEKFYSCCEPYDILFVNLTRKTLFTYVNL 248
 Qy 250 VPCVLLASMLLGLFTLPDSDSEKLSGLVLTLLSLTVFLNMVAETMPATSDAVPLLGTVFN 309

Db 249 IPCWGISFLSVLYFLPSDGEKISLCISILLSITVFFLLABRIIPPTSLTVPLGKYL 308
 QY 310 CIMFVASSVSTILINHYHRHADTHEMDWIRCVFLYWLFWLWLRMSRPGSATTPPPAR 369
 Db 309 FTMLVTLSSVVVIAVLNFRFVTHRMAPVQRLFIQLPKLCTIERP--KKEEPEED 366
 QY 370 VPP-----PPDLR-LRERSKSL-----L 387
 Db 367 QPPEVLTDVYVHLFPDVKFVNYDSKPSGDIYGPALPASHRFDLAAGGISAHCFAEPPL 426
 QY 388 ANVLDI---DDPRHFQAO---QPOCCRYRGGEENGAGLAHRS-----CFGVDY--ELS 434
 Db 427 PSSLPPLGADDDLFSPSLGNDISPGCC-----PAAAAAADAADLSPFFKPYAREME 478
 QY 435 LILKEIRVITDQWRKDEDEDADISRDNKFAMVVDRLCLIFTFTIATLVALLSAPHI 493
 Db 479 KTIEGSRFIAQHVKNKDKFESVEDWKYVAMVLDRLMFLWIFATAICVVGTAIIILQAPSL 537
 RESULT 13
 ACH2_RAT STANDARD; PRT; 511 AA.
 AC P12389; 008952;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Neuronal acetylcholine receptor protein, alpha-2 chain precursor.
 GN CHR2A2 OR ACRA2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;
 RX MEDLINE=88178113; PubMed=2832952;
 RA Wada K., Ballivet M., Boulter J., Connolly J.G., Wada E.,
 RA Deneris E.S., Swanson L.W., Heinemann S.F., Patrick J.;
 RT "Functional expression of a new pharmacological subtype of brain
 RT nicotinic acetylcholine receptor.";
 RL Science 240:330-334 (1988).
 RN [2]
 RP REVISIONS.
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;
 RA Boulter J.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: After binding acetylcholine, the AChR responds by an
 CC extensive change in conformation that affects all subunits and
 CC leads to opening of an ion-conducting channel across the plasma
 CC membrane.
 CC -!- SUBUNIT: Neuronal AChR seems to be composed of two different types
 CC of subunits: alpha and non-alpha (beta). Alpha-2 subunit can be
 CC combined to beta-2 or beta-4 to give rise to functional receptors.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL; L10077; AAB60900.1; -;
 DR EMBL; M20297; AAA40664.1; -;
 DR EMBL; M20292; AAA40664.1; JOINED.
 DR EMBL; M20293; AAA40664.1; JOINED.
 DR EMBL; M20294; AAA40664.1; JOINED.
 DR EMBL; M20295; AAA40664.1; JOINED.
 DR EMBL; M20296; AAA40664.1; JOINED.
 DR PIR; A40110; A40110.
 DR InterPro; IPR006029; Neu_channel_memb.

DR InterPro; IPR006029; Neu_chan_LBD.
 DR InterPro; IPR006201; Neur_channel.
 DR Pfam; PF02931; Neur_chan_LBD; 1.
 DR Pfam; PF02932; Neur_chan_memb; 1.
 DR PRINTS; PR00252; NRIONCHANNEL.
 DR TIGRFAMS; TIGR00860; LIC; 1.
 DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
 KW Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
 KW Transmembrane; Multigene family.
 FT SIGNAL 1 27 POTENTIAL.
 FT CHAIN 28 511 NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,
 FT DOMAIN 28 241 ALPHA-2 CHAIN.
 FT TRANSMEM 242 266 EXTRACELLULAR.
 FT TRANSMEM 274 292 POTENTIAL.
 FT TRANSMEM 308 329 POTENTIAL.
 FT DOMAIN 330 484 CYTOPLASMIC.
 FT TRANSMEM 485 503 POTENTIAL.
 FT DOMAIN 391 402 POLY-GLU.
 FT DISULFID 160 174 BY SIMILARITY.
 FT DISULFID 224 225 ASSOCIATED WITH RECEPTOR ACTIVATION
 FT CARBOHYD 56 56 (BY SIMILARITY).
 FT CARBOHYD 106 106 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 212 212 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CONFLICT 494 494 C -> S (IN REF. 1; AAA40664).
 SQ SEQUENCE 511 AA; 58611 MW; 3824E83BB01D613B CRC64;
 Query Match 35.7%; Score 951; DB 1; Length 511;
 Best Local Similarity 39.4%; Pred No. 3.3e-72;
 Matches 198; Conservative 90; Mismatches 167; Indels 48; Gaps 8;
 QY 16 LLLLLCLMPRGARCGYHEKRLHLHLHDHYNLVRPVVNESDPLQLSFLGLTLMQIIDVDE 75
 Db 19 LLLVPAVLQQGSHT-HAEDRLFKHLFGGYNWARVPNTSDVIVRFGLSLAQLIDVDE 77
 QY 76 KQOLLITNLWLKLEWNNLNRNTSDFGKXKDLRPPHRLKPKDVLNYSADGEPSTYP 135
 Db 78 KNQMTTNVWLKQEWNDYKLRWDPAEFGNVTSLRVPSEMIWIPDVLNYSADGEPSTYP 137
 QY 136 TNVVRNNGSLYVPPGIFKSTCKIDITWFPDDQRCMKFGSWTVGYQLDLQDQEGG 195
 Db 138 TXAHLLFTGTWVPPAIYKSSCSIDVTFPDDQRCMKFGSWTVYDKAKIDLE-QMERT 196
 QY 196 GDISSFVTNGEWELIGVPGKRNIEYNCPEYDITFAVIRRKTYLYFFNLIVPCVLI 255
 Db 197 VDLKDYWESGEWALLNATGTYSKKYDCCAEIYPDVYVYFVIRRLPLFTVNLIIIFCLLI 256
 QY 256 ASWALLGFTLPPDSGEKLSGLVTILLSLTVLNMVAETMPATSDAVPLGLTYFNCIMFWY 315
 Db 257 SCLTVLVFLYLPSECEKLTCLISVLLSLTVFLLITEIIPSTSLVLIPLIGEYLLFTMIV 316
 QY 316 ASSVSTIILINHYHRHADTHEMSDWIRCVFLYWLFWLWLRMSRPGSATTPPARVPPPD 375
 Db 317 TLSIVITVFLNVHRSPTNMPWVRVALLGRVPRWLMNRP-----LPPWELHGSFD 371
 QY 376 LELRERSKSLANVLDDDFRHFQAOQPOCCRYRGGEEN---GAGLAHRSFCFGVDY- 431
 Db 372 LKL--SPSVHWTETNMDAGERETEETEEEEE-----DENICVAGL-PDSSMGVLKG 420
 QY 432 -----ELSLILKEIRVITDQWRKDEDEDADISRDNKFAM 465
 Db 421 HGLHLRAMEPTKTPSQASEILLSPQIKALEGVHYIADRLSRSEDADSSVKEDWKYVAM 480
 QY 466 VVDRCLLIIFTFTIATLAVLL 488
 Db 481 VVDRFLWFLIIVCFLGTIGLFL 503
 RESULT 14
 ACH3_BOVIN
 ID ACH3_BOVIN STANDARD; PRT; 495 AA.
 AC Q07263;

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:30:03 ; Search time 31.1579 Seconds

(without alignments)
5022.709 Million cell updates/sec

Title: US-09-303-232-4

Perfect score: 2665

Sequence: 1 MCGRRRRSLAAPAGLLLL.....LFTIATLAVLLSAPHIMWS 496

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25: *
1: sp archaea: *
2: sp bacteria: *
3: sp fungi: *
4: sp human: *
5: sp invertebrate: *
6: sp mammal: *
7: sp mhc: *
8: sp organelle: *
9: sp phage: *
10: sp plant: *
11: sp rodent: *
12: sp virus: *
13: sp vertebrate: *
14: sp unclassified: *
15: sp virus: *
16: sp bacteriaph: *
17: sp archaea: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2665	100.0	496	5	Q9XZI3
2	2064.5	77.5	807	5	Q87V5
3	2047	76.8	542	5	Q86MN7
4	2046.5	76.8	545	5	Q9VW19
5	1803.5	67.7	501	5	Q9XZI4
6	1795.5	67.4	494	5	Q87S2
7	1792.5	67.3	494	5	Q87S1
8	1791.5	67.2	494	5	Q86MN8
9	1790.5	67.2	494	5	Q87S3
10	1774	66.6	509	5	Q87S0
11	1769	66.4	523	5	Q87R9
12	1484.5	55.7	554	5	Q9VLT9
13	1371	51.4	525	5	Q81PE2
14	1272.5	47.7	509	13	Q800C7
15	1253	47.0	555	13	Q7T2U0
16	1251	46.9	570	5	Q9VJT9

17	1242	46.6	502	11	Q9JHD6
18	1239.5	46.5	486	13	Q7T2S0
19	1235	46.3	502	6	Q866A2
20	1219.5	45.8	511	13	Q03481
21	1207	45.3	513	13	Q7T2R9
22	1166.5	43.8	554	13	Q7T2T9
23	1150.5	43.2	474	13	Q7T2U1
24	1141	42.8	461	5	P91197
25	1113	41.8	480	5	Q81932
26	1106.5	41.5	335	5	Q9NKD1
27	1058.5	39.7	554	5	Q62083
28	1055.5	39.6	542	5	Q18556
29	994	37.3	523	5	Q46128
30	975.5	36.6	537	5	Q9U941
31	972	36.5	515	5	Q46133
32	962	36.1	537	5	Q8MUR0
33	961	36.1	505	4	Q8U777
34	960	36.0	536	5	Q8T0Y9
35	956.5	35.9	509	5	Q8MM21
36	951	35.7	536	5	Q8T9S0
37	949.5	35.6	509	5	Q9NFX8
38	947	35.5	568	5	Q9NFR5
39	942.5	35.4	552	5	P91765
40	940.5	35.3	512	11	Q91X60
41	937.5	35.2	545	5	Q96631
42	932	35.0	499	11	Q8VHH6
43	929	34.9	540	5	O46134
44	928	34.8	499	11	Q8R4G9
45	928	34.8	504	11	Q8BV44

ALIGNMENTS

RESULT 1

Q9XZI3	PRELIMINARY;	PRT;	496 AA.
ID	Q9XZI3		
AC	Q9XZI3;		
DT	01-NOV-1999 (TREMBLrel. 12, Created)		
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)		
DE	01-JUN-2003 (TREMBLrel. 24, Last annotation update)		
DE	Putative nicotinic acetylcholine receptor alpha 7-1 subunit.		
OS	Heliothis virescens (Noctuid moth) (Owlet moth).		
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;		
OC	Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;		
OC	Noctuidae; Heliothinae; Heliothis.		
OX	NCBI_TaxID=7102;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Schulte T., Oellers N., Adamczewski M.;		
RT	"Putative alpha subunits of insect nicotinic acetylcholine receptors		
RT	more similar to vertebrate alpha 7 subunits and C. elegans Ce21 than		
RT	to other insect nicotinic acetylcholine receptor alpha subunits.";		
RL	Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.		
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).		
CC	-1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.		
DR	EMBL; AF143846; AAD32697.1; -		
DR	GO; GO:0016021; C:integral to membrane; IEA.		
DR	GO; GO:0005230; F:extracellular ligand-gated ion channel acti...; IEA.		
DR	GO; GO:0005216; F:ion channel activity; IEA.		
DR	GO; GO:0030594; F:neurotransmitter receptor activity; IEA.		
DR	GO; GO:0006811; P:ion transport; IEA.		
DR	GO; GO:0007269; P:synaptic transmission; IEA.		
DR	InterPro; IPR006201; Neur channel.		
DR	InterPro; IPR006202; Neur_chan_LBD		
DR	InterPro; IPR006029; Neur_chan_LBD_memb.		
DR	Pfam; PF02931; Neur_chan_LBD; 1.		
DR	Pfam; PF02932; Neur_chan_memb; 1.		
DR	PRINTS; PR00252; NRIONCHANNEL.		
DR	TIGRFAMs; TIGR00860; LIC; 1.		
DR	PROSITE; PS00236; NEUROTR ION CHANNEL; 1.		
KW	Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;		
KW	Transmembrane.		

```

SQ SEQUENCE 496 AA; 56347 MW; 8032FED8515A6210 CRC64;
Query Match 100.0%; Score 2665; DB 5; Length 496;
Best Local Similarity 100.0%; Pred. No. 9.4e-251;
Matches 496; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGRARRSHLAAPAGALLLCLLPARGCGYHEKRLHLLDHYNVLERPVVNESDPLQ 60
DB 1 MGRARRSHLAAPAGALLLCLLPARGCGYHEKRLHLLDHYNVLERPVVNESDPLQ 60
QY 61 LSFGLTLMQIIDVDEKNQLITNIWLEWDMNLRWNTSDFGVKDLRVPVHRLWKPDV 120
DB 61 LSFGLTLMQIIDVDEKNQLITNIWLEWDMNLRWNTSDFGVKDLRVPVHRLWKPDV 120
QY 121 LMYNSADEGSDSYPTNVVVRNNGSCLYVPPGIFKSTCKIDITWFPDDQRCCKMFGSWT 180
DB 121 LMYNSADEGSDSYPTNVVVRNNGSCLYVPPGIFKSTCKIDITWFPDDQRCCKMFGSWT 180
QY 181 YDGYQLDLQDEGGDISFVINGEWELIGVPGKNEIYVNCPEPIYDITFAVVIIRK 240
DB 181 YDGYQLDLQDEGGDISFVINGEWELIGVPGKNEIYVNCPEPIYDITFAVVIIRK 240
QY 241 TLYYFENLIVPCVLIIASMLLGTLPDPSGKLSLGVITLLSTVFLNMVAETMPATSDA 300
DB 241 TLYYFENLIVPCVLIIASMLLGTLPDPSGKLSLGVITLLSTVFLNMVAETMPATSDA 300
QY 301 VPLLGTTFNCIMFWASSVSTILINHYHRRHADTHEMSDWIRCVFLYVWVLRMSRPG 360
DB 301 VPLLGTTFNCIMFWASSVSTILINHYHRRHADTHEMSDWIRCVFLYVWVLRMSRPG 360
QY 361 SATTPPPARVPPPPDLERSSKSLANVLIDDDFRHPQAOQPCCRYYRGGEEGAG 420
DB 361 SATTPPPARVPPPPDLERSSKSLANVLIDDDFRHPQAOQPCCRYYRGGEEGAG 420
QY 421 LAHSCFGVYELSLILKEIRVITDQMKDDEDADISRDWKFAAMVVDRLCLIFTFTI 480
DB 421 LAHSCFGVYELSLILKEIRVITDQMKDDEDADISRDWKFAAMVVDRLCLIFTFTI 480
QY 481 IATLAVLLSAPHIMVS 496
DB 481 IATLAVLLSAPHIMVS 496

RESULT 2
Q87TV5 PRELIMINARY; PRT; 807 AA.
ID Q87TV5 AC Q87TV5
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nicotinic acetylcholine receptor Dalpha5 subunit.
GN NACR-ALPHA-34E OR NACRALPHA-34E OR BG:DS05899.4 OR BG:DS05899.5 OR
CG4498 OR CG16878 OR CG32975.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
MEDLINE=21969411; PubMed=11973307;
RA Grauso M., Reenan R.A., Culetto E., Sattelle D.B.;
RT "Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
RT Dalphas5, Dalphas6 and Dalphas7, in Drosophila melanogaster Identify a
RT New and Highly Conserved target of Adenosine Deaminase Acting on RNA-
RT Mediated A-to-I Pre-mRNA Editing.";
RL Genetics 160:1519-1533(2002).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR EMBL; AF272778; AAM13390.1; -.
DR FlyBase; FBgn0028875; nAcR-alpha-34E.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. .; IEA.
```


OS	Heliothis virescens (Noctuid moth) (Owlet moth).
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC	Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;
OC	Noctuidae; Heliothinae; Heliothis.
OX	NCBI_TaxID=7102;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Schulte T., Oellers N., Adamczewski M.;
RT	"Putative alpha subunits of insect nicotinic acetylcholine receptors
RT	more similar to vertebrate alpha 7 subunits and C. elegans Ce21 than
RT	to other insect nicotinic acetylcholine receptor alpha subunits."
CC	Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RL	- - SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC	- - SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR	EMBL; AF143847; AAD32698.1; -
DR	GO; GO:0016021; C:integral to membrane; IEA.
DR	GO; GO:0005230; F:extracellular ligand-gated ion channel acti. . . ; IEA.
DR	GO; GO:0005216; F:ion channel activity; IEA.
DR	GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
DR	GO; GO:0006811; P:ion transport; IEA.
DR	GO; GO:0007268; P:synaptic transmission; IEA.
DR	InterPro; IPR006201; Neur channel.
DR	InterPro; IPR006202; Neur_chan_LBD.
DR	InterPro; IPR006029; Neur_channel_memb.
DR	Pfam; PF02931; Neur_chan_LBD; 1.
DR	Pfam; PF02932; Neur_chan_memb; 1.
DR	PRINTS; PR00252; NRIONCHANNEL.
DR	TIGRFAMS; TIGR00860; LIC; 1.
DR	PROSITE; PS00236; NEUROTR_ION CHANNEL; 1.
KW	Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW	Transmembrane.
SQ	SEQUENCE 501 AA; 56704 MW; 43CB0DC3960C78AB CRC64;
QY	Query Match 67.7%; Score 1803.5; DB 5; Length 501;
Db	Best Local Similarity 68.7%; Pred. No. 8.9e-157;
QY	Matches 347; Conservative 53; Mismatches 78; Indels 27; Gaps 7;
QY	12 APAGLLLLCLLWPRGARGYHEKRLHLLHLLHYNLPERPVNESDPLQSFGLTLMQII 71
Db	2 APMLAALLALLPVSEQ-GPHEKRLHLLHLLHYNLPERPVNESDPLQSFGLTLMQII 60
QY	72 DVEKNOLLITNWLKLEWDMNLRNWTSDFGVKDLRVPVPHRLKPDVLMYNSADEGFD 131
Db	61 DVEKNOLLITNWLKLEWDMNLRNWTSDFGVKDLRVPVPHRLKPDVLMYNSADEGFD 120
QY	132 STYPTNVVNRNNGSLYVPPGIFKSTCKIDITWFPDDQRCMKFGSWTYDGYQLDLQ 191
Db	121 GYQTNVVRSGGSLYVPPGIFKSTCKIDITWFPDDQRCMKFGSWTYDGYQLDLQ 180
QY	192 DEGGDISPVINGEWELIGVPGKRNIEYVNCPEPIDITFAVIRKRLTYFFNLVLP 251
Db	181 DEAGDLSDFITNGEWELIGVPGKRNIEYVNCPEPIDITFAVIRKRLTYFFNLVLP 240
QY	252 CVLIASMALLGFTLPDPSGKLSGLVITLLSLVFLNMVAETPATSDAVPLIGTYFNCI 311
Db	241 CVLISSMALLGFTLPDPSGKLSGLVITLLSLVFLNMVAETPATSDAVPLIGTYFNCI 300
QY	312 MFMVASSVSTILILNVHHRHADTHEMSDWIRCVFLWLPWLRMSRPGSATTPPPARV- 370
Db	301 MFMVASSVSTILILNVHHRHADTHEMSDWIRCVFLWLPWLRMSRPGSATTPPPARV- 360
QY	371 PPPDLELRSSKSLANVLIDDDFRH-----PQAQQ-----PQCCRY---YRGG--- 414
Db	361 TMRELELKERSSKSLANVLIDDDFRHGGPPPNSTAGTNGNPGSCISFRTDFRFSVR 420
QY	415 -----EENGAGLAHSCFVGVDYELSLILKEIRVITDQMKDEDEDADISDMKFAAMVVDRL 470
Db	421 PSTMEDVGGGLGSH-----HRELHLILRELOFITARMKKADEAEELISDWKFAAMVVD 475
QY	471 CLLIIFTLTIIATIAVLISAPHIMV 495
Db	476 CLFVFTLTIIATIAVLISAPHIMV 500

```

RESULT 6
Q8T7S2
ID Q8T7S2 PRELIMINARY; PRT; 494 AA.
AC Q8T7S2;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Nicotinic acetylcholine receptor Dalpha6 subunit variant type II.
GN NACR-ALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21969411; PubMed=11973307;
RA Grauso M., Reenan R.A., Culetto E., Sattelle D.B.;
RT "Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
RT Dalpha5, Dalpha6 and Dalpha7, in Drosophila melanogaster Identify a
RT New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-
RT Mediated A-to-I Pre-mRNA Editing."
RL Genetics 160:1519-1533 (2002).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR EMBL; AF321446; AAM13393.1; -.
DR FlyBase; FBgn0032151; nacr-alpha-30D.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. .; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR GO; GO:0007268; P:synaptic transmission; IEA.
DR InterPro; IPR006201; Neur_chan.
DR InterPro; IPR006202; Neur_chan_LBD.
DR InterPro; IPR006029; Neu_chan_LBD.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_LBD; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMs; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
DR GlycoProtet; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
SQ SEQUENCE 494 AA; 56048 MW; 6EE711810EDE7BBB CRC64;

Query Match 67.4%; Score 1795.5; DB 5; Length 494;
Best Local Similarity 68.1%; Pred. No. 5.3e-166;
Matches 342; Conservative 54; Mismatches 81; Indels 25; Gaps 5;

QY 8 SHLAAPAGLLLLCLLWPRGARC-GVHEKRLHLLDHNVLPRVPPVNSDPLQLSFGLT 66
Db 3 SLPASLSLFLVLIIFAIKESCOGHEKRLHLLSTNTLPRVANESEPLEVFKGLT 62
QY 67 LMQIIDVDEKQQLITNIWLKLEWMDNMLRWNTSDFGVKDLRVPVPHRLKPDVLMYNSA 126
Db 63 LQOIIDVDEKQQLITNLWLSLEWMDNMLRWNETYGVKDLRITNKLKPDVLMYNSA 122
QY 127 DEGFSTYTNVVRNGSLYVPPGIFKSTCKIDITWFFDDQRCMKFGSWTYDGYQL 186
Db 123 DEGFDTYHTNIVVRHGSCLYVPPGIFKSTCKMDITWFFDDQHCMEKFGSWTYDGNQL 182
QY 187 DLQIQDEGGDTSFVTNGEWELIGVPGKRNELTYNCCPEYIDITFAVVIRKTLTYFF 246
Db 183 DIVINSEDDGDLSDFTNGEWLLAMPKGNKTIIVYACCPPEYVDITFTIQRRTLYYFF 242
QY 247 NLIVPCVLIASMALLGFTLPDPSGKLSLGVILLISLTVFLNVAETMPATSDAVELIGT 306
Db 243 NLIVPCVLISSMALLGFTLPDPSGKLTGLVILLISLTVFLNVAESMTTSDAVPLIGT 302
QY 307 YFNCIMFWASSVSTILINVHHRHADTHEMSDWIRCVFLYWLWLVLMRSPGSAITPP 366
Db 303 YFNCIMFWASSVLTIVVLYVNHRTADITHPPPIKSVFLQWLPWLIRMGPRGKITRK 362

```

```

Db 123 DEFGDTYHTNVVKGSGCLYVPGIFKSTCKMDITWFFDDOHCMEKFGSWTYDGNQL 182
QY 187 DLOQDEGGDISFVTNGEWELIGVPGKNEIYVNCPPPYDITFVAVIRKTLTYFF 246
Db 183 DLVNSGDDGLSDFTNGEWYLLAMPKKNITVYACCPYVDITFTIQRRTLYFF 242
QY 247 NLIVPCVLIASMALLGFTLPDGSKEKSLGVTLISLTVFLNVAETMPATSDAVPLLGT 306
Db 243 NLIVPCVLISSMALLGFTLPDGSKEKSLGVTLISLTVFLNVAETMPATSDAVPLLGT 302
QY 307 YFCINPMWASSVSTILLNHYHRRHADTHEMSDWIRCVFLYLPWVLRMSRPGSATTPP 366
Db 303 YFCINPMWASSVVLTVVLNHYHRTADIHMPWPKSVFLQWLPWLRMGRPKITRK 362
QY 367 PARVP-PPDLELRERSKSLANVLDDDDFRHPQAOQCRCYRGGGANGAGLAA-- 423
Db 363 TILLSNMKELELERSKSLANVLDDDDFRHTIS------GSQTAGSSASF 411
QY 424 IFTLFTTIATLAVLLSAPHIMV 495
Db 472 VFTLFTTIATVLLSAPHIIV 493

```

```

RESULT 8
Q86MN8
ID Q86MN8 PRELIMINARY; PRT; 494 AA.
AC Q86MN8
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Nicotinic acetylcholine receptor subunit Dalpha6 precursor.
GN NACRALPHA-30D.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
CX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Millar N.S.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ554209; CAD86935.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. .; IEA.
DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR InterPro; IPR006201; Neur_chan_LBD.
DR InterPro; IPR006202; Neur_chan_LBD.
DR InterPro; IPR006029; Neur_chan_LBD.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_LBD; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMs; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
KW SIGNAL.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 494 NICOTINIC ACETYLCHOLINE RECEPTOR SUBUNIT
FT CHAIN 27 494 DALPHA6.
SQ SEQUENCE 494 AA; 55980 MW; 4DFC572139587070 CRC64;

```

```

Query Match 67.2%; Score 1791.5; DB 5; Length 494;
Best Local Similarity 68.1%; Pred. No. 1.3e-165;
Matches 342; Conservative 54; Mismatches 81; Indels 25; Gaps 5;
QY 8 SHLAAPAGLLLCLLWPRGARC-CYHEKRLHLHLLHYNVLRPVNPNESDPLQSLFGIT 66
Db 3 SPLPASLSFLVLLIFLAIKESQGPHEKRLHLHLLSYNTLTPRVAESEPLEVKFGIT 62

```

```

QY 67 LMQIIDVDEKNQLLITNIWLKLEWNDNLRWNTSDFGGVKDLRVPPHRLWKPDVLMNSA 126
Db 63 LQIIDVDEKNQLLITNIWLKLEWNDNLRWNTSDFGGVKDLRVPPHRLWKPDVLMNSA 122
QY 127 DEFGDTYHTNVVKGSGCLYVPGIFKSTCKMDITWFFDDOHCMEKFGSWTYDGNQL 186
Db 123 DEFGDTYHTNVVKGSGCLYVPGIFKSTCKMDITWFFDDOHCMEKFGSWTYDGNQL 182
QY 187 DLOQDEGGDISFVTNGEWELIGVPGKNEIYVNCPPPYDITFVAVIRKTLTYFF 246
Db 183 DLVNSGDDGLSDFTNGEWYLLAMPKKNITVYACCPYVDITFTIQRRTLYFF 242
QY 247 NLIVPCVLIASMALLGFTLPDGSKEKSLGVTLISLTVFLNVAETMPATSDAVPLLGT 306
Db 243 NLIVPCVLISSMALLGFTLPDGSKEKSLGVTLISLTVFLNVAETMPATSDAVPLLGT 302
QY 307 YFCINPMWASSVSTILLNHYHRRHADTHEMSDWIRCVFLYLPWVLRMSRPGSATTPP 366
Db 303 YFCINPMWASSVVLTVVLNHYHRTADIHMPWPKSVFLQWLPWLRMGRPKITRK 362
QY 367 PARVP-PPDLELRERSKSLANVLDDDDFRHPQAOQCRCYRGGGANGAGLAA-- 423
Db 363 TILLSNMKELELERSKSLANVLDDDDFRHTIS------GSQTAGSSASF 411
QY 424 IFTLFTTIATLAVLLSAPHIMV 495
Db 472 VFTLFTTIATVLLSAPHIIV 493

```

```

RESULT 9
Q87S3
ID Q87S3 PRELIMINARY; PRT; 494 AA.
AC Q87S3;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nicotinic acetylcholine receptor Dalpha6 subunit variant type I.
GN NACRALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
CX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Grauso M., Reenan R.A., Culetto E., Sattelle D.B.;
RT "Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes, Dalpha5, Dalpha6 and Dalpha7, in Drosophila melanogaster Identify a New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-Mediated A-to-I Pre-mRNA Editing."
RT Mediated A-to-I Pre-mRNA Editing."
RL Genetics 160:1519-1533(2002).
CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR EMBL; AF321445; AAML3392.1; -.
DR FlyBase; FBgn0032151; nAChR-alpha-30D.
DR GO; GO:0016020; C:integral to membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. .; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR GO; GO:0007268; P:synaptic transmission; IEA.
DR InterPro; IPR006201; Neur_chan_LBD.
DR InterPro; IPR006202; Neur_chan_LBD.
DR InterPro; IPR006029; Neur_chan_LBD.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_LBD; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMs; TIGR00860; LIC; 1.

```

DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
 KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
 SQ SEQUENCE 494 AA; 56095 MW; B46BEDA63A92942 CRC64;
 Query Match 67.2%; Score 1790.5; DB 5; Length 494;
 Best Local Similarity 67.9%; Pred. No. 1.6e-165;
 Matches 341; Conservative 53; Mismatches 83; Indels 25; Gaps 5;
 QY 8 SHLAAPAGLLLLCLLWPRGARC-GYHEKRLHLHLHDHYNVLRPVVNESDPLQLSFGLT 66
 Db 3 SPLPASLSFLVLLIFLAIKESCGQPKHRLNHLSTNTLRPVANSEPLEVFGLT 62
 QY 67 LMQIIDVDEKQNLITNWLKLEWMDNLRWNTSDFGVKKDLRVP 126
 Db 63 LQQLIDVDEKQNLITNWLKLEWMDNLRWNTSDFGVKKDLRIT 122
 QY 127 DEGFSTYPTNVVRNNGSCLYVPPGIFKSTCKIDITWFPDDQRCMKFGSWTYDGYQL 186
 Db 123 DEGFDTYHTNIVVKHNGSCLYVPPGIFKSTCKIDITWFPDDQRCMKFGSWTYDGNQL 182
 QY 187 DLQLODEGGDISFVNGEWELIGVPGKNEIYVNCBPPIYDITFVAVIRKTLIYFF 246
 Db 183 DLVNSDEGDLSDFTNGEWYLLAMPKKNITVYACCPYVDITFTIQRRTLIYFF 242
 QY 247 NLIVPCVLIASMALLGFTLPDPSGKLSLGVTLISLTVPLNVAETMATSDAVPLGT 306
 Db 243 NLIVPCVLISSMALLGFTLPDPSGKLSLGVTLISLTVPLNVAETLPQVSDAIPGT 302
 QY 307 YFNCIMFVASSVSTILINLHHRHADTHEMSDWIRCVFLYVLPVLRMSRPGSATTP 366
 Db 303 YFNCIMFVASSVVLTVVNLVYHRTADIEHMPPIKSVFLQWLPVLRMGRGRKTIK 362
 QY 367 PARVP-PPDLELRSSKSLANVLDDDDFRHPQAQQPCCRYPQCCRY 423
 Db 363 TILLSNRMKELEKERSKSLANVLDDDDFRHTIS-----GSQTAIGSSASF 411
 QY 424 -----HSCFGVDY-ELSLILKEIRVITDQMRKDEADISRDMKFAAMVVDRLCLT 473
 Db 412 GRPTTVEEHHTAGCNHKLHLILKELOFTATMRKADDEAELIGDKWFAAMVVDRLCLT 471
 QY 474 IFTLFTIATLAVLLSAPHIMV 495
 Db 472 VFTLFTIATVTLSSAPHIV 493
 RESULT 10
 Q8T7S0 PRELIMINARY; PRT; 509 AA.
 AC Q8T7S0;
 DT 01-JUN-2002 (TREMELrel. 21, Created)
 DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
 DE Nicotinic acetylcholine receptor Dalpha6 subunit variant type IV.
 GN NACR-ALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21969411; PubMed=11973307;
 RA Grauso M., Reenan R.A., Culetto E., Sattelle D.B.;
 RT "Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
 RT Dalpha5, Dalpha6 and Dalpha7, in Drosophila melanogaster Identify a
 RT New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-
 RT Mediated A-to-I Pre-mRNA Editing.";
 RL Genetics 160:1519-1533 (2002).
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
 DR EMBL; AF321448; AAM13395.1; -;
 DR FlyBase; FBgn0032151; nAcR-alpha-30D.

DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti...; IEA.
 DR GO; GO:0005216; F:ion channel activity; IEA.
 DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
 DR GO; GO:0006811; P:ion transport; IEA.
 DR GO; GO:0007268; P:synaptic transmission; IEA.
 DR InterPro; IPR006201; Neur_chan.
 DR InterPro; IPR006202; Neur_chan_LBD.
 DR InterPro; IPR006029; Neu_Channel_memb.
 DR Pfam; PF02931; Neur_chan_LBD; 1.
 DR Pfam; PF02932; Neur_chan_memb; 1.
 DR PRINTS; PR00252; NEIONCHANNEL.
 DR TIGRfams; TIGR00860; LIC; 1.
 DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
 KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
 KW Transmembrane.
 SQ SEQUENCE 509 AA; 57887 MW; BE8D8E0198B0C2BD CRC64;
 Query Match 66.6%; Score 1774; DB 5; Length 509;
 Best Local Similarity 66.0%; Pred. No. 6.9e-164;
 Matches 341; Conservative 54; Mismatches 82; Indels 40; Gaps 6;
 QY 8 SHLAAPAGLLLLCLLWPRGARC-GYHEKRLHLHLHDHYNVLRPVVNESDPLQLSFGLT 66
 Db 3 SPLPASLSFLVLLIFLAIKESCGQPKHRLNHLSTNTLRPVANSEPLEVFGLT 62
 QY 67 LMQIIDV-----DEKQLITNWLKLEWMDNLRWNTSDFGVKKDLRVP 111
 Db 63 LQQLIDVDEKQNLITNWLKLEWMDNLRWNTSDFGVKKDLRIT 122
 QY 112 PHRLKPDVLMNSADEGDSYPTNVVRNNGSCLYVPPGIFKSTCKIDITWFPDDQ 171
 Db 123 PNKLKPDVLMNSADEGDSYPTNVVRNNGSCLYVPPGIFKSTCKIDITWFPDDQ 182
 QY 172 CEMKFGSWTYDGYQLDLOLQDEGGDISFVNGEWELIGVPGKNEIYVNCBPPIYDI 231
 Db 183 CEMKFGSWTYDGNQLDLVNSDEGDLSDFTNGEWYLLAMPKKNITVYACCPYVDI 242
 QY 232 TFAVIRKTLIYFFNLIVPCVLIASMALLGFTLPDPSGKLSLGVTLISLTVFLNVA 291
 Db 243 TFTIQRRTLIYFFNLIVPCVLISSMALLGFTLPDPSGKLSLGVTLISLTVFLNVA 302
 QY 292 ETMPATSDAVPLIGYFNCIMFVASSVSTILINLHHRHADTHEMSDWIRCVFYWLP 351
 Db 303 ETLQPVSDAIPLLGTYFNCIMFVASSVVLTVVNLVYHRTADIEHMPPIKSVFLQWLP 362
 QY 352 WYLRMSRPGSATTPPARVP-PPDLELRSSKSLANVLDDDDFRHPQAQQPCCRYP 410
 Db 363 WYLRMGRGRKTIKTRKILLSNRMKELEKERSKSLANVLDDDDFRHTIS----- 414
 QY 411 YRGGEENGAGLAA-----HSCFGVDY-ELSLILKEIRVITDQMRKDEADISR 458
 Db 415 ---GSQTAIGSSASGRPTTVEEHHTAGCNHKLHLILKELOFTATMRKADDEAELIG 471
 QY 459 DMKFAAMVVDRLCLIFLFTIATLAVLLSAPHIMV 495
 Db 472 DMKFAAMVVDRLCLIFLFTIATVTLSSAPHIV 508
 RESULT 11
 Q8T7R9 PRELIMINARY; PRT; 523 AA.
 ID Q8T7R9;
 AC Q8T7R9;
 DT 01-JUN-2002 (TREMELrel. 21, Created)
 DT 01-JUN-2002 (TREMELrel. 21, Last sequence update)
 DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
 DE Nicotinic acetylcholine receptor Dalpha6 subunit variant type V.
 GN NACR-ALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

```

RN Q9VL79
RP SEQUENCE FROM N.A.
RX MEDLINE=21969411; PubMed=11973307;
RA Grauso M., Reenan R.A., Culetto E., Sattelle D.B.;
RT "Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
RT Dalphas, Dalphas and Dalphas7, in Drosophila melanogaster identify a
RT New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-
RT Mediated A-to-I Pre-mRNA Editing.";
RL Genetics 160:1519-1533(2002).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR EMBL; AF321449; AM13396.1; -
DR FlyBase; FBgn0032151; nAChR-alpha-30D.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel activity; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR GO; GO:0007268; P:synaptic transmission; IEA.
DR InterPro; IPR006201; Neur channel.
DR InterPro; IPR006202; Neur channel.
DR InterPro; IPR006029; Neur_channel_memb.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_memb; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRfams; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
SQ SEQUENCE 523 AA; 59110 MW; 1C200AF4F87F841 CRC64;

Query Match 66.4%; Score 1769; DB 5; Length 523;
Best Local Similarity 64.4%; Pred. No. 2.2e-163;
Matches 342; Conservative 53; Mismatches 82; Indels 54; Gaps 6;

Qy 8 SHLAAPAGLLLLCLLWPGARC-GYHEKLLHLLDHYNVLPRPVNVEDPQLSPGLT 66
Db 3 SPLPASLSLVLLIFLAIKESQCGPHEKLLHLLSYNTLPRPVANSEPLEVKEFLT 62

Qy 67 LMOIIDVDENKOLLITNMLKLENDNMLRWNTSDFGVGKDLRPPRLMKPDVLMVNSA 126
Db 63 LQIIDVDENKOLLITNMLKLENDNMLRWNTSDFGVGKDLRPPRLMKPDVLMVNSA 122

Qy 127 DEGFDSYTPNVRNNGSLYVPPGIFKSTCKIDITWFFDDQRCMKFGSWTYDGYQL 186
Db 123 DEGFDTGTYHTNIVKNGSLYVPPGIFKSTCKWDITWFFDDQRCMKFGSWTYDGNQL 182

Qy 187 DLQIQDEGGDISFVTVNGENELGVPGKNEIYNNCCPEPYDITFAVIRKTLTYFF 246
Db 183 DLVINSDDGDLSDFITNGEWYLLAMPKKNITVYACCPPEYVDITFIQIRRTLYYFF 242

Qy 247 NLIVPCVLIASMALLGFTLPDPSGKLSLGVTLISLTFTVLNVAETMPATSDAV---- 301
Db 243 NLIVPCVLISSMALLGFTLPDPSGKLTGLVTLLSLTFTVLNVAESMPPTSDAVPLIG 302

Qy 302 -----PLLGTYFCNIMFMVSSVSTIILNHYHHRHADTHE 337
Db 303 TILLSLTFTVLNVAETLPQVSDAIFLLGTYFCNIMFMVSSVLTIVVLNHYHRTADIE 362

Qy 338 MSWIRCVFLYLPWLVRMRGPGATTPPARVP-PPDLELRSSKSLIANVLIDDD 396
Db 363 MPPIKSVFLQWLVRMRGPGKTRKRTILLNRMKELEKELKRSKSLIANVLIDDD 422

Qy 397 FRHPQAQPCCRVYRGEENGAGLAA-----HSCFGVDY-ELSLILKEIRVIT 444
Db 423 FRHTIS-----GSQTAGSSASFRPTVEHHHTAIGCNKDLHLILKELOFIT 471

Qy 445 DQMKEDDEDADISRDWKAFAAMVDRLCLIIFTFTIATIALVLSAPHIMV 495
Db 472 ARNRKADDEAEELIGDWKAFAAMVDVDFCLIVFTFTIATVTLVLSAPHIIV 522
```

RESULT 12

```

Q9VL79
ID Q9VL79; PRELIMINARY; PRT; 554 AA.
AC Q9VL79;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE C94128 protein.
GN NACR-ALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Calle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton R.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abrell J.F., Agbayani A., An H.-U., Andrews-Frankkoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foslter C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Ananides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Faragas V., Park S., Patel S., Pfeiffer B., Scheeler F.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Swirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
```


QY 375 DLRLRRSSKLLANVLIDDDDR-----HPQAQPO-----CCRYRGGREN-----CAGL 421
Db 370 DLNLSPGVAOSTNGNLIIYI--GFGMTIHH-YATSPDSGWICSLVATGEDVLLPGAQA 426
QY 422 AAHSCFGV-DYELSLIILKEIRVITDQMRKODEDADISRDWKFAAMVVDRLCLIIFTLFTI 480
Db 427 SSVSSSGFGETELSKILDEVRIYSKRFRDQDEEDTVCNEWKFAASVIDRLCLMAFSLFTI 486
QY 481 IATLAVLLSAPHIM 494
Db 487 LCTIGILMSAPNFV 500

RESULT 15

Q7T2U0 PRELIMINARY; PRT; 555 AA.
AC Q7T2U0
DT 01-OCT-2003 (T-EMBLrel. 25, Created)
DT 01-OCT-2003 (T-EMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Nicotinic acetylcholine receptor alpha 7b subunit (Fragment).
GN A7B.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RA Jones A.K., Elgar G., Sattelle D.B.;
RT "The nicotinic acetylcholine receptor gene family of the pufferfish,
RT Fugu rubripes";
RL Genomics 0:0-0(2003).
DR EMBL; AY298752; AAP57216.1; -.
KW Receptor.
FT NON_TER
SQ SEQUENCE 555 AA; 61010 MW; B042A3E4430A2B7C CRC64;

Query Match 47.0%; Score 1253; DB 13; Length 555;
Best Local Similarity 44.7%; Pred. No. 4.7e-113;
Matches 246; Conservative 86; Mismatches 126; Indels 92; Gaps 9;
QY 31 GYHEKRLHLLHLYNVLRRPVVNESDPLQLSGLTLMQIIDVDKXQLLIINILKLEW 90
Db 4 GPEQRTLLKXLLKDYNRMRPVANDSQPLTVVFTLSLIQIMDVDEKNQILTTNILRMSG 63
QY 91 NDMLRWNTSDFGVKDLRVPVPHRLKPDVLMYNSADEGFDSTYPTNVVVRNNGSLYVP 150
Db 64 FDHLYQWQSEHPGVKNLFTTDQIWTDPILLYNSADDDFDSTFTKNLVNSSGYAEYQP 123
QY 151 PGIFKSTKIDITWPFDDQRCMKFGSWTYDGYQLDQLQDEGGDISFVTNGEWELI 210
Db 124 PGIFMSTCNVDVRFPPFDIQCRLKFGSWTYDGLLDLQWNE---ADISGYMANGEWDLI 180
QY 211 GVPGRKRNEYCCPEPIDITFAVVRKTYFFENLIVPCVLIASMALLGFTLPDPSG 270
Db 181 GVPGRTRNEFYOCKEPYPAVTFVVAIRRTIYALNLLIPCVLLSSMTLLIFVLPA DSG 240
QY 271 EKLSLGVTLISLTVFLNVAETMPATSDAVPLGLTYFNCIMFMVASSVVTILINLHH 330
Db 241 EKISLITVLLSVLVAEIMPATSDVPLIGQYFASIMIIVGMSSVAVVVLQYHH 300
QY 331 RHADTHEMSDNRVFLYLPWVLRMSRFGSATP--PP-----ARVPPPP 374
Db 301 HDENGNGMKPVQLVLLQWVAVFLMRKPGKXDSRPPCAPHLRRCSSGSGSIPNPP 360
QY 375 DLLEL-----RERSKSLIANVL-----DIDDFRHPQAQ-- 404
Db 361 DHALHPLHPQGLAPLQPGHLAGQPHVHAQSSANNNGNLVILFGQSVDSAGLPEPIQRN 420
QY 405 -----PQCCRY-----YRGEENGAGLAHS 425

Db 421 NISTGPPRVAGSPPHLPSPQFCSSPPPPASNMMDTGCPSTVSSGGFGG---GGGLGGCS 477
QY 426 CFGV-DYELSLIILKEIRVITDQMRKODEDADISRDWKFAAMVVDRLCLIIFTLFTIATL 484
Db 478 ASAVGDPQLHAILLEEVRFVADRFREQDEVSGAADQWKFAGAVIDRLCLVAFSVFNIICTI 537
QY 485 AVLLSAPHIM 494
Db 538 SILMSAPNFV 547

Search completed: May 7, 2004, 11:38:38
Job time : 33.1579 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 8, 2004, 11:50:34 ; Search time 3463.3 Seconds
(without alignments)
4276.744 Million cell updates/sec

Title: US-09-303-232-4

Perfect score: 2665

Sequence: 1 MCGRRARRSHAPAGLLLL.....LFTTIATLAVLLGAPHIMVS 496

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh
-O=/cgn2.1/USPTO.spool/US09303232/runat_07052004.101110.23893/app_query.fasta_1.2261
-DB=EST -QFMT=fastap -SUFFIX=rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09303232 @CGN 1.1 4577 @runat_07052004.101110.23893 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DRV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST.*
1: em_estba.*
2: em_esthum.*
3: em_estin.*
4: em_estmu.*
5: em_estov.*
6: em_estpl.*
7: em_estro.*
8: em_htc.*
9: gb_est1.*
10: gb_est2.*
11: gb_hic.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: em_estfun.*
16: em_estom.*
17: em_gss_hum.*
18: em_gss_inv.*
19: em_gss_pln.*
20: em_gss_vrt.*
21: em_gss_fun.*
22: em_gss_mam.*
23: em_gss_mus.*
24: em_gss_pro.*
25: em_gss_fod.*
26: em_gss_phg.*
27: em_gss_vrl.*
28: gb_gss1.*

29: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
C					
1	53.8	1433	885	12	BG632919 GH16126.3
2	36.0	960	2296	14	CD013901
3	35.4	942.5	1436	29	AY402873 Homo sapi
4	34.8	928	1864	11	AK053497 Mus muscu
5	34.8	928	2916	11	AK051730 Mus muscu
6	34.8	928	3126	11	AK080415 Mus muscu
7	34.8	927.5	1201	9	AL530299 AL530299
8	34.3	914.5	4037	11	AK049722 Mus muscu
9	34.3	914.5	4046	11	AK051742 Mus muscu
10	34.3	913.5	1436	29	AY402875 Mus muscu
11	34.0	905	1454	29	AY402878 Mus muscu
12	33.6	895.5	1442	29	AY402876 Mus muscu
13	33.5	892.5	922	13	BUS15857
14	33.4	889	1374	29	AY406230
15	33.4	889	1374	29	AY406232
16	33.4	889	4290	11	AK029177 Mus muscu
17	33.1	881.5	2010	11	AK080475 Mus muscu
18	32.9	877.5	1442	29	AY402877 Mus muscu
19	32.6	869	2940	11	AK034228 Mus muscu
20	32.6	869	3230	11	AK083157 Mus muscu
21	32.5	866.5	1781	14	CD013889
22	32.0	853	4589	11	AK030464 Mus muscu
23	31.7	845.5	1374	29	AY406231 Pan trogl
24	31.4	835.5	3483	11	AK081254 Mus muscu
25	31.2	832.5	1436	29	AY402874 Mus muscu
26	30.1	801.5	1751	14	CD013905
27	30.1	801.5	2110	14	CD013902
28	30.1	801.5	2110	14	CD013903
29	29.2	779	1466	14	CD013891
30	29.2	777	2074	11	AK080554 Mus muscu
31	28.9	771.5	2257	11	AK017571 Mus muscu
32	28.9	771	2513	11	AK033068 Mus muscu
33	28.9	769	1034	13	EX403124
34	28.6	762.5	908	13	BUI49265
35	27.5	733.5	607	9	AI292581
36	27.0	720.5	1596	14	CD013888
37	26.8	715.5	615	14	CB149460
38	26.4	704	1603	29	AY411327
39	26.3	700.5	2948	11	AK081344 Mus muscu
40	25.8	686.5	833	14	CB245337
41	25.6	682.5	1101	29	CNS0001F
42	25.4	677.5	2782	11	AK047974
43	25.4	676.5	1007	29	CNS000HJU
44	25.3	674.5	902	29	AY407186
45	25.2	672	902	29	AY407184 Homo sapi

ALIGNMENTS

RESULT 1

BG632919/c

LOCUS

DEFINITION

BG632919 885 bp mRNA linear EST 23-APR-2001
GH16126.3prime GH Drosophila melanogaster head p072 Drosophila
melanogaster cDNA clone GH16126.3 similar to CG4128: FBan0004128
'ion channel' located on: 2L 30D1-30E1; 04/10/2001, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Drosophila melanogaster (fruit fly)

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 885)

REFERENCE
AUTHORS Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
Lewis,S. and Rubin,G.M.
BDGP/HMI Drosophila EST Project
JOURNAL Unpublished (2001)
COMMENT Other_ESTs: GH16126.5prime
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu
Based upon the presence of a XhoI site followed by a run of 14 or
more T residues at the beginning of the sequence, this clone was
polyadenylated. The resulting Poly-T sequence has been removed. hit
genomic AEO03511: arm:X (18792641,19136447)
estimated-cyto:18A3-18C6: 04/10/2001
Plate: GH.161 row: C column: 2
High quality sequence stop: 784.
Location/Qualifiers
1..885
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="GH16126"
/sex="male and female"
/dev_stage="adult"
/lab_host="DH5 - alpha"
/clone_lib="GH Drosophila melanogaster head pOT2"
/note="Organ: head; Vector: pOT2; Site:1: EcoRI; Site 2:
XhoI; Sized fractionated cDNAs were directly ligated into
pOT2. Plasmid cDNA library."

ORIGIN

Alignment Scores:
Pred. No.: 5,12e-122 Length: 885
Score: 1433.00 Matches: 260
Percent Similarity: 97.18% Conservatve: 16
Best Local Similarity: 91.55% Mismatches: 6
Query Match: 53.77% Indels: 2
DB: 12 Gaps: 1

US-09-303-232-4 (1-496) x BG632919 (1-885)

Qy 65 LeuThrLeuMetGlnIleAspValaspGluLysAsnGlnLeuLeuIleThrAsnIle 84
Db 885 CTAACACTCATGCAGATTATCGATGTCGACGAGAGAATCAACTGCTTATAACGAATAT 826

Qy 85 TrpLeuLysLeuGluTrpAsnAspMetAsnLeuArgTrpAsnThrSerAspPheGlyGly 104
Db 825 TGGCTCAAAATGGAAATGGAACGATGATGAATCTTCATGGAATTCGATGAGTTTCGGTGGT 766

Qy 105 ValLysAspLeuArgValProHisArgLeuTrpLysProAspValLeuMetTyrAsn 124
Db 765 GTGCGGATCTCGCAATTCGCCCATCGCTATGGAACCGGATGATGATGATGATCAAC 706

Qy 125 SerAlaAspGluGlyPheAspSerThrTrpProThrAsnValValArgAsnGly 144
Db 705 AGTGCAGCAGAGGGCTTCGATGGAACGATGACCCCAATGGTGGTTCGCAATTAATGGG 646

Qy 145 SerCysLeuTrpValProGlyIlePheLysSerThrCysLysIleAspIleThrTrp 164
Db 645 AGCTGTCGTACGTACCCAGGTATATTAAGTCAACGTTAGATGACATACGTTGG 586

Qy 165 PheProPheAspGlnArgCysGluMetLysPheGlySerTrpThrTrpThrAspGlyTyr 184
Db 585 TTTCCATTTCGACGATCAGATGTAATGAAATTTGTTGTTGTCGACCTACGATGGGTT 526

Qy 185 GlnLeuAspLeuGlnLeuGlnAspGluGlyGlyAspIleSerSerPheValThrAsn 204
Db 525 CAGTTGGACCTGAGTTGCGAGACGAGCTGGTGGCGACATTTCTAGCTTTTATAACCAAT 466

Qy 205 GlyGluTrpGluLeuIleGlyValProGlyLysArgAsnGluIleTyrTyrAsnCysCys 224
Db 465 GCGCAATGGGACTTGTAGGTGTCGCCGTAACCAAAATCAATCTACTATAATTCCTGC 406

Qy 225 ProGluProTyrIleAspIleThrPheAlaValAlaIleArgArgLysThrLeuTyrTyr 244
Db 405 CCAGAACCTTATATGACATAACATTCGCCATTTTGTATAGGCGCAAAAGCTTGTACTAT 346

Qy 245 PhePheAsnLeuIleValProCysValLeuIleAlaSerMetAlaLeuLeuGlyPheThr 264
Db 345 TTTTTCATCTGATTTGTCGCTGCGTACTGATCGCTCCATGCGCTAGGTTTAC 286

Qy 265 LeuProProAspSerGlyGluLysLeuSerLeuGlyValThrIleLeuLeuSerLeuThr 284
Db 285 CTGCCACCAAGATTCTCGTGAAAAGCTTTCGTTGGAGTTTACATTTTATTCGCTTAC 226

Qy 285 ValPheLeuAsnMetValAlaGluThrMetProAlaThrSerAspAlaValProLeuLeu 304
Db 225 GTCTTCTCAACATGTCGCGCAACATGCCGCGACCTCCGATGGGTACCGCTGCTC 166

Qy 305 GlyThrTyrPheAsnCysIleMetPheMetValAlaSerSerValValSerThrIleLeu 324
Db 165 GGAACCTTATTCATTTGCATTATGTTTATGGTGGCTCATCATGTTGTCAACCATCTT 106

Qy 325 IleLeuAsnTyrHisHisArgHisAlaAspThrHisGluMetSerAspTrpIleArgCys 344
Db 105 GTCTCAATTTATCATCATAGAAATCCAGATACGATGAAATGATGATGATGATGATG 52

Qy 345 ValPheLeuTyr 348
Db 51 ATAGTTTATAT 40

RESULT 2

LOCUS CD013901 2296 bp mRNA linear EST 21-OCT-2003
DEFINITION 90134548 Single gene library Homo sapiens cDNA, mRNA sequence.
ACCESSION CD013901
VERSION CD013901.1 GI:37777431
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2296)
AUTHORS Jin,P., Fu,G.K., Wilson,A.D., Yang,J., Chien,D., Hawkins,P.R.,
Au-Young,J. and Stuve,L.L.
TITLE PCR isolation and cloning of novel splice variant mRNAs from known
drug target genes
JOURNAL Unpublished (2003)
COMMENT Contact: Jin, P.
3160 Porter Drive, Palo Alto, CA 94304, USA
Tel: 650 621 8639
Fax: 650 621 8965
Email: pjin@incyte.com.
Location/Qualifiers
1..2296
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="Single gene library"
/note="Vector: pDrive Cloning Vector; RT-PCR was performed
using gene-specific primers flanking the open-reading
frame. PCR products were subcloned into pDrive Cloning
Vector and sequenced completely using M13 forward and
reverse primers. Sequencing gaps were closed by
re-sequencing using primers flanking the gapped areas."

ORIGIN

Alignment Scores:
Pred. No.: 1.31e-77 Length: 2296
Score: 960.00 Matches: 199
Percent Similarity: 56.45% Conservatve: 90

		Best Local Similarity:	38.87%	Mismatches:	171
		Query Match:	36.02%	Indels:	52
		DB:	14	Gaps:	8
US-09-303-232-4 (1-496) x CD013901 (1-2296)					
QY	4	ArgAlaArgSerHisLeuAlaProAlaGlyLeuLeuLeuCysLeuLeu	23		
Db	278	AGGGCTCCTGGAGACCACTCTCTCCAGTCCACCAGCATTCG	322		
QY	24	TripProArgGlyAlaArgCysGlyTyrrHisGluLysArgLeuLeuHisHisLeuLeuAsp	43		
Db	323	---CCGAGGAGCGTCGCATACCGAGACTGAGGCCGCTCTTCAAACACCTCTTCGG	379		
QY	44	HisTyrrAsnValLeuGluArgProValValAsnGluSerAspProLeuGlnLeuSerPhe	63		
Db	380	GGGTACAACCGCTGGCGCGCGCGGTGCCAACACTTCAGAGCTGGTGATTTGGCTTT	439		
QY	64	GlyLeuThrLeuMetGlnIleleAspValAspGluLysAsnGlnLeuLeuThrAsn	83		
Db	440	GGACTCTCCATCGCTCAGCTCATCGATGTGGATGAGAACAACAAATGATGACCAAC	499		
QY	84	IleTripLeuLysLeuGlnTrpAsnAspMetAsnLeuArgTrpAsnThrSerAspPheGly	103		
Db	500	GTCTGCTTAACAGAGGTGGAGGACACTCAAACTGCGGTGAACCCCGCTGATTTGGC	559		
QY	104	GlyValLysAspLeuArgValProHisArgLeuTrpLysProAspValLeuMetTyr	123		
Db	560	AACATCACATCTCTCAGGCTCCCTTCTGAGATGATCTGGATCCCCGACATTGTCTCTAC	619		
QY	124	AsnSerAlaAspGluGlyPheAspSerThrTyrPrOthrAsnValValAsnAsn	143		
Db	620	AACATGAGATGGGGAGTTTCAGTAGTACGCCAATGACCAATGTTCTCCACG	679		
QY	144	GlySerCysLeuTyrValProGlyPheLysSerThrCysLysIleAspIleThr	163		
Db	680	GGCATGTGCATGGGTGCCCCGCCACACTCAAGAGCTCTCGAGCATCGAGTCAAC	739		
QY	164	TriPheProPheAspAspGlnArgCysGluMetLysPheGlySerThrTyrAspGly	183		
Db	740	TTCTTCCCCTTCGACGACGAGAACTGCAAGATGAAGTTGGCTCTGCACTATGACAA	799		
QY	184	TyrGlnLeuAspGlnLeuGlnAspGluGlyGlyAspIleSerSerPheValThr	203		
Db	800	GCCAAGATGACCTGGAG---CAGATGGAGCAGACTGTGGACCTGAAGCACTACTGGG	856		
QY	204	AsnGlyGluTrpGluLeuGlyValProGlyLysArgAsnGluLeuTyrTyrrAsnCys	223		
Db	857	AGCGCGAGTGGCCATCGTCATAATGCCAGCGGCACCTCAACAGCAAGATGACAGTC	916		
QY	224	CysProGluProTyrIleAspIleThrPheAlaValValIleArgArgLysThrLeuTyr	243		
Db	917	TSGCGCGAGATCTACCCGACGCTACCTACGCTTCGTTCATCGCGGCTGCCCTCTTC	976		
QY	244	TyrPhePheAsnLeuLeuValProCysValLeuIleAlaSerMetAlaLeuGlyPhe	263		
Db	977	TACACATCAACTCATCATCCCTGCCTGCTCATCTCTGCTCCTGCTGCTGCTCTTC	1036		
QY	264	ThrLeuProProAspSerGlyLysLeuSerLeuGlyValThrIleLeuLeuSerLeu	283		
Db	1037	TACCTCCCTCCGACGCGCGAGAGATFACGCTGTCATTTGGTCTGCTGCTCACTC	1096		
QY	284	ThrValPheLeuAsnMetValAlaGluThrMetProAlaThrSerAspAlaValProLeu	303		
Db	1097	ACGGTCTTCTGCTGCTCATCTAGATCATCCCGTCCACTCGCTGGTCACTCCGCTC	1156		
QY	304	LeuGlyThrTyrrPheAsnCysIleMetPheMetValAlaSerSerValValSerThrIle	323		
Db	1157	ATCGCGGAGTACTGCTGTGTACCATGATCTTCGTTCACCTGTCTCATCCCGTCTC	1216		
QY	324	LeuIleLeuAsnTyrrHisHisArgHisAlaAspThrHisGluMetSerAspTrpIleArg	343		
Db	1217	TTCGTCTCAATGTGCACACCGCTCCCGCCACACCAACCATGCCCCACTGGTGGCG	1276		

```

gene      <1..>1436
/ gene="CHRNA3"
/ locus_tag="HCML369"

ORIGIN
Alignment Scores:
Pred. No.: 2,75e-76      Length: 1436
Score: 942.50           Matches: 194
Percent Similarity: 58.25% Conservative: 92
Best Local Similarity: 39.51% Mismatches: 153
Query Match: 35.37%      Indels: 53
DB: 29                  Gaps: 6

US-09-303-232-4 (1-496) x AY402873 (1-1436)
Qy 28 AlaArgCysGlyTyrHisGluLysArgLeuLeuHisLeuLeuAspHisTyrAsnVal 47
Db 3 GCCAGGCGCTCAGAGGCTGAGCACCGTCTATTGAGCGGCTGTTGAGGATTACAATGAG 62
Qy 48 LeuGluArgProValValAsnGluSerAspProLeuGlnLeuSerPheGlyLeuLeu 67
Db 63 ATCATCGGCTGTGGCCACAGTGTCTGACCCAGTCAATCCATTTGAGGTGCTCATG 122
Qy 68 MetGlnIleLeuAspValAspGluLysAsnGlnLeuLeuLeuLeuThrAsnIleTrpLeuLys 87
Db 123 TCTCAGCTGTGAAGGTGAAGTAAGTAACACAGATCATGGAGACCAACCTGTGGCTCAAG 182
Qy 88 LeuGluTrpAsnAspMetAsnLeuLeuArgTyrAsnThrSerAspPheGlyValLysAsp 107
Db 183 CAATCTGGAACTGACTCAAGCTGAAGTGAACCCCTCTGACTATGTGGGGCAGAGTTC 242
Qy 108 LeuArgValProProHisArgLeuTrpLysProAspValLeuMetTyrAsnSerAlaAsp 127
Db 243 ATCGGTCTCCTGCACAGAGATCTGGAAGCCAGACATTTGCTGTATACAAATGCTGTT 302
Qy 128 GluGlyPheAspSerThrTyrProThrAsnValValValArgAsnAsnGlySerCysLeu 147
Db 303 GGGGATTTCCAGGTGGGAGCAACAGACCAAGCCCTTACTCAAGTACACTGGGAGGTGACT 362
Qy 148 TyrValProProGlyIlePheLysSerThrCysLysIleAspIleThrTrpPheProPhe 167
Db 363 TGGATCTCTCCGGCCATCTTTAAGAGCTCTCTGTAATAATCGACGTGACTTCCCGTTT 422
Qy 168 AspAspGlnArgCysGluMetLysPheGlySerThrTyrAspGlyTyrGlnLeuAsp 187
Db 423 GATTACCAAACTGTACCATGAAGTTCGGTTCCTGGTCTCAGTAAAGGCGAAATCGAT 482
Qy 188 LeuGlnLeuAspGluGlyGlyAspIleSerSerPheValThrAsnGlyGluTrp 207
Db 483 CTGGTCTGTATCGGC---TCTTCATGAACCTCAAGGACTATTGGGAGAGCGCGAGTGG 539
Qy 208 GluLeuIleGlyValProGlyLysArgAsnGluIleTyrTyrAsnCysCysProGluPro 227
Db 540 GCCATCATCAAGCCCGGCTCAAAACAGACATCAAGTACACTGCTCGAGGAGTAC 599
Qy 228 TyrIleAspIleThrPheAlaValIleArgLysThrLeuTyrTyrPhePheAsn 247
Db 600 TACCCGACATCATACTGCTGTACATCGGGCGCTGCTGCTGTCTACACCATCAAC 659
Qy 248 LeuIleValProCysValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProPro 267
Db 660 CTCATCATCCCTGCTGCTCATCTCTCTCTCACTGTGCTGCTCTTACCTGCCCTCC 719
Qy 268 AspSerGlyGluLysLeuSerLeuGlyValThrIleLeuLeuSerLeuValPheLeu 287
Db 720 GACTGCGGTGAGAGGAGGACCTGTGATTTCTGCTCTCTCTCCCTGACCGGTGTTCTC 779
Qy 288 AsnMetValAlaGluThrMetProAlaThrSerAspAlaValProLeuLeuGlyThrTyr 307
Db 780 CTGGTGTACTGAGACCATCCCTTCCACCTCGTGTGCTATCCCTGATTGGAGAGTAC 839
Qy 308 PheAsnCysIleMetPheMetValAlaSerSerValValSerThrIleLeuIleLeuAsn 327

```

```

Db 840 CTCCTGTTCCACCATGATTTTGTAAACCTTGTCCATCGTCATCACCGTCTTGTGCTCAAC 899
Qy 328 TyrHisHisArgHisAlaAspThrHisGluMetSerAspTrpIleArgCysValPheLeu 347
Db 900 GTGCACCTACAGAACCCCGACACACACAAATGCCCTCATGGGTGAAGACTGTAATCTTG 959
Qy 348 TyrTrpLeuProTrpValLeuArgMetSerArgPro-----GlySerAlaThr 363
Db 960 AACCTGCTCCAGGGTCACTGTTTCATGACAGGCCCAACAGCAACGAGGCGCAACGCTCAG 1019
Qy 364 ThrPro-ProProAlaArgValProProProProProProProProProProProProPro 383
Db 1020 AAGCCGAGGCGCTCTACGTTGCGC-----AGCTC 1049
Qy 383 rLysSerLeuLeuAlaAsnValLeuAspIleAspAspPheArgHisProGlnAlaG1 403
Db 1050 TCAAAATCTGAATGCTTCAGCC-----GGCA 1076
Qy 403 nGlnProGlnCysCysArgTyrTyrArgGlyGlyGluGluAsnGlyAlaGlyLeuAla1 423
Db 1077 GAGTCCAAAGGC-TGCAAGGAGGGCTACCCCTGCCAGGAGGGATGTGTGTACTGCCA 1135
Qy 423 aHis-----SerCysPh 427
Db 1136 CCACCGCAGGATAAAATCTCAATTTTCAGTGTCTAACCTCAGGAGAAGCTCTAGTCTGA 1195
Qy 427 eGlyValAspTyrGluLeuSerLeu-----IleLeuLy 438
Db 1196 ATCTGTGTAGTGTGTGCTGCTGCTCTCTCTGCTTTGTCCACGAAATCAAGAGCATCA 1255
Qy 438 sGluIleArgValIleThrAspGlnMetArgLysAspAspGluAspAlaAspIleSerAr 458
Db 1256 AAGTGTCAAGTATATGCTGAAATATGAAGACACAAATGAAGCCAAAGAGATTCAAGA 1315
Qy 458 gAspTrpLysPheAlaAlaMetValValAspArgLeuCysLeuIleIlePheThrLeuPh 478
Db 1316 TGATTGAAAGTATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1375
Qy 478 eThrIleAlaThrLeuAlaValLeuLeu 488
Db 1376 GTGCATTTAGGGACACGAGGATTGTTCTCG 1406

RESULT 4
AK053497
LOCUS
DEFINITION
Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched library, clone E130103E14 product: NEURONAL NICOTINIC ACETYLCHOLINE RECEPTOR ALPHA 3 SUBUNIT, full insert sequence.
ACCESSION
VERSION AK053497.1 GI:26343494
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
Mus musculus
REFERENCE
1 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)
2 20499374
PUBMED
11042159
AUTHORS
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

```


Qy	273	LeuSerLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAenMetValAlaGlu	292
Db	892	GTACGCTCTGCATCTCGTGCTCTCTCCTCGAGGCTCTTCTCCTCGTGATCACCGAG	951
Qy	293	ThrMetProAlaThrSerAspAlaValProLeuLeuGlyThrTyrPheAenCysIleMet	312
Db	952	ACCATCCTTCCACCTCACTGGTGCATCCCTTGATCGGGAGTAGTACCTCTCTTCACTATG	1011
Qy	313	PheMetValAlaSerSerValValSerThrIleLeuLeuLeuAenTyrHisHisArgHis	332
Db	1012	ATTTTTCACCTTGTCCATCGTCATCACAGTCTTTGTGTCACAGTGCCAGCAACT	1071
Qy	333	AlaAspThrHisGluMetSerAspTyrIleArgCysValPheLeuTyrTyrLeuProTyr	352
Db	1072	CCGACACACACACAGATGCCACTTGGTCAAGCTGTGTTTGAACCTTCTCCCCAGG	1131
Qy	353	ValLeuArgMetSerArgProGlySerAlaThrThrProProProAlaArgValProPro	372
Db	1132	GTCAATGTTTATGACTAGGCCA-----ACCAGCACCGAGGAGAGCGCCCAAG	1179
Qy	373	ProProAspLeuGluLeuArgGluArgSerSerLysSerLeuLeuAlaAsnValLeuAsp	392
Db	1180	ACGAGGAACCTTACGGTGTGAGCTCTCAACCTGAACTGCTTCAGCCGTGCAGACTCC	1239
Qy	393	IleAspAspAspPheArgHisProGlnAlaGlnGlnProGlnCys-----CysArgTyr	410
Db	1240	AAAAGCTCAAGGAGGCTACCCC---TGCCAAGATGGGACATGTGGCTATTGCCACAC	1296
Qy	411	TyrArgGlyGlyGluGluAsn---GlyAlaGlyLeuAla---AlaHisSerCysPheGly	428
Db	1297	CGTAGGGTAAAAATCTCAATTTTCAGTGCACCACTCAAGAAGAGCTCCAGTTCTGAGTCT	1356
Qy	429	ValAspTyrGluLeuSerLeu-----IleLeuLysGlu	439
Db	1357	GTTCATGCTGTGTGTGCTCTCTCTGTCTGTACACAGAAATCAAGAAGCCATCAAGT	1416
Qy	440	IleArgValIleThrAspGlnMetArgLysAspAspGluAspAlaAspIleSerArgAsp	459
Db	1417	GTGAAGTATATTGCTGAAATATGAAGACACAGAATGAGCCAAAGAGATTCAAGATGAT	1476
Qy	460	TrpLysPheAlaAlaMetValValAspArgLeuCysLeuIleIlePheThrLeuPheThr	479
Db	1477	TGGAAGTATGTGTGCATGTGATTCATGTCATTTTCTCTGGTTCCTCATCTGCTGTGTC	1536
Qy	480	IleIleAlaThrLeuAlaValLeuLeu	488
Db	1537	ATTTAGGAACATGCAGGATTATTCTG	1563
RESULT 5			
AK051730			
LOCUS	AK051730	2916 bp	mRNA linear HTC 20-SEP-2003
DEFINITION	Mus musculus 12 days embryo spinal ganglion cDNA, RIKEN full-length enriched library, clone:D130068A06 product:NEURONAL NICOTINIC ACETYLCHOLINE RECEPTOR ALPHA 3 SUBUNIT, full insert sequence.		
ACCESSION	AK051730		
VERSION	AK051730.1	GI:26342155	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS	1 Carninci, P. and Hayashizaki, Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE	2		
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
20499374			
11042159			
3			
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuami, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.			
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer			
Genome Res. 10 (11), 1757-1771 (2000)			
20530913			
11076861			
4			
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.			
Functional annotation of a full-length mouse cDNA collection			
Nature 409, 685-690 (2001)			
5			
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.			
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs			
Nature 420, 563-573 (2002)			
6 (bases 1 to 2916)			
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saigo, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, K., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.			
Direct Submission			
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)			
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.			
Please visit our web site for further details.			
URL: http://genome.gsc.riken.go.jp/			
URL: http://fantom.gsc.riken.go.jp/			
Location/Qualifiers			
1. .2916			
/organism="Mus musculus"			
/mol_type="mRNA"			
/strain="C57BL/6J"			
/db_xref="FANTOM_DB:D130068A06"			
/db_xref="MGI:2420066"			
/db_xref="taxon:10090"			
/clone="D130068A06"			
/tissue type="spinal ganglion"			
/clone_lib="RIKEN full-length enriched mouse cDNA library"			
/dev_stage="12 days embryo"			
170. .1669			
/note="unnamed protein product; NEURONAL NICOTINIC ACETYLCHOLINE RECEPTOR ALPHA 3 SUBUNIT (SPTR)AAL58471, evidence: FASTV, 99.8%ID, 100%length, match=1497"			
putative"			
/codon_start=1			
CDS			

|||||
592 TTTCCCTTGTGTCAGCAGCTGCAAACTGAAGTTTGGTCTCTGCTCTACGGAGCTGG 651
|||||
185 GlnLeuAspLeuGlnLeuGlnAspGluGlyGlyAspIleSerPheValThrAsn 204
|||||
652 TCCTTGGATCTGCAGATGAGGAG-----CCAGATATFCAGTGGCTATATCCCAAT 702
|||||
205 GlyGluTrpGluLeuLeuGlyValProGlyLysArgAsnGluIleTyrTyrAsnCysCys 224
|||||
703 GGAGATGGGACCTAGTGGGATCCCGGCAAGAGAGTGAAGAGTCTATCAGTKCTGC 762
|||||
225 ProGluProTyrIleAspIleThrPheAlaValValIleArgArgLysThrLeuTyrTyr 244
|||||
763 AAAGAGCCCTACCCCGATGTCACTTCACAGTGACCATGCGCGCAGGAGCTCTACTAT 822
|||||
245 PhePheAsnLeuIleValProCysValLeuIleAlaSerMetAlaLeuLeuGlyPheThr 264
|||||
823 GGCCTCAACCTGCTGATCCCTGTGTCTCATCTCCGCCCTCGCCCTGCTGCTCTCTG 882
|||||
265 LeuProProAspSer-GlyGluLysLeuSerLeu-GlyValThrIleLeuLeuSerLeu 284
|||||
883 CTTCCTGCAGATTCGGGGGAGAGATTCCCTKGGGATPAACAGTATTACTCTCTTA 942
|||||
284 hrValPheLeuAsnMetValAlaGluThrMetProAlaThrSerAspAlaValProLeu 304
|||||
943 CCGCTTTCATGCTGCTGCTGATCATCGCCGCAACATCCGATCCGTTACCATTTGA 1002
|||||
304 euGlyThrTyrPheAsnCysIleMetPheMetValAlaSerSerValValSerThrIle 324
|||||
1003 WAGCCAGTCTTCGCCGAGCACCATGATCATCGTGGGCTTTGG-GTGGTTGKRMGGTGA 1061
|||||
324 euLeuLeuAsnTyrHis 329
|||||
1062 TCGGTGCTGCAGTACCAC 1078
|||||

RESULT 8
AK049722
LOCUS
DEFINITION Mus musculus 12 days embryo spinal cord cDNA, RIKEN full-length enriched library, clone:CS30044P16 product:cholinergic receptor, nicotinic, beta polypeptide 2 (neuronal), full insert sequence.
ACCESSION AK049722.1 GI:26340459
VERSION
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
METH. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
2
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
MEDLINE 11042159
PUBMED
3
Shibata,K., Itoh,M., Aizawa,K., Nagaoaka,S., Sasaki,N., Carninci,P., Konno,H., Akiyama,J., Nishi,K., Kitsuami,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,K., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
CDS
ORIGIN

20530913
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 4037)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W., Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T., Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M., Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M., Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/
Location/Qualifiers
1. .4037
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:CS30044P16"
/db_xref="MGI:2417229"
/db_xref="taxon:10090"
/clone="CS30044P16"
/tissue_type="spinal cord"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="12 days embryo"
200. .1705
/note="unnamed protein product; cholinergic receptor, nicotinic, beta polypeptide 2 (neuronal) (MGI:MGI-87891, GB|NM_009602, evidence: BLASTN, 99%, match=1498) putative"
/codon_start=1
/protein_id="BAC33893.1"
/db_xref="GI:26340460"
/translations="MARCSNMFSLFGLLWLCGVLGTDTTERLVHLLDPSRYNK LIPATNGSELVTQVMVSLAQLISVHEREQIMTNVMTQEWEDYRLTWKPFDFNM KVLRLSKHTWLEDVLYNNADGKYEVSFVSNVSDGSIFFWLPPAIYSACKIEVK HFFPDQNCMTKFRSWTDRIEIDLVLKSDVASIDDDFTSGFMDWIDIALPGRNENPD STVYDITDFIIRKKPLYNILNIPCVLITSLAILVFLPSDCGKXKTLCTISVLLAL TVFLISIKTVPPTSLDVLPGKYLMTVLVTFISVTVCLNVHRSPTHTMAP VKVFLKPLTLFLQQRHRCQRQLRLRRRREGAGTLFFREGPAADPCTCFVY PASMQGLAGAFQAEPAAGGSRMGPCSGLREAVDGVRFIADHRSDEDDQSVREDW KYAVMWIDRLFLMIFVFCVFGTIGMLQLPQNYTATTFLHSDHSPSSK"


```

Db      1082 CTGAGCTACCGCTGGTGGGAAGTACCTCATGTTCCACGANGGTGTAGTACCTTCTCC 1141
QY      319 ValValSerThrIleLeuLeuLeuAsnTyrHisArgHisAlaAspThrHisGluMet 338
Db      1142 ATGCTCACTAGCGTGTGTGCTCAATGTGCACACCGTTCGCTACCGACACCATG 1201
QY      339 SerAspTrpIleArgCysValPheLeuTyrTrpLeuProTrpValLeuArgMetSerArg 358
Db      1202 GCOCCTGGGTCAAGGTGGTCTTCCTGGAGAAGCTGCCACCTCTCTTCCTGCGACGAG 1261
QY      359 ProGlySerAlaThrThrProProAlaArgValProProProProProProProProPro 378
Db      1262 CCAGCCACCGCTGT-----GCAGT-----CAACCCCTGGCGCTTG 1297
QY      379 ArgGluArgSerSerIysSerLeuLeuAlaAsnValLeuAspIleAspAspPheArg 398
Db      1298 CGAAGGCGCCAGCGGAACGTGAGGGGCGACGACACTA-----TTCTTCGCG 1345
QY      399 HisProGlnMlaGlnGlnProGlnCysCysArgTyrTyrArgGlyGlyGluGluAsn--- 417
Db      1346 GAAGGTCTCTGACGTGACCATGACCTGCTTTGTCAATCCTGCATCAATGCGAGGCGTTG 1405
QY      418 ---GlyValGlyLeuAlaAlaHisSerCysPheGlyValAspTyrGluLeu----- 433
Db      1406 GCTGGGGCTTTCAGGCTGAGCCGCTGACCGCGTGGCGCTGATGGGCGCCCATGC 1465
QY      434 SerLeuIleLeuIysGlu-----IleArgValIleThrAspGlnMetArgLys 449
Db      1466 AGCTGTGGCTCCGGGAAGCTGTGACGGTGTGATTCATTGGGACCATATGCGAAGT 1525
QY      450 AspAspGluAspAlaAspIleSerArgAspTrpIysPheAlaAlaMetValValAspArg 469
Db      1526 GAAGATGATACACAGAGTGTGAGGAGGATGGAAATACGTTGCCATGGTGTGATCGACCGC 1585
QY      470 LeuCysLeuIleIlePheThrLeuPheThrIleAlaThrLeuAlaValLeuLeu 488
Db      1586 CTGTCTCTGGATCTTGTCTGTCTGTCTGTCTGTGGACCATGGCAGTTCCTG 1642

RESULT 10
LOCUS   AY402875
DEFINITION Mus musculus CHRNA3 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY402875
VERSION   AY402875.1 GI:39758858
KEYWORDS  GSS.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferrlera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL  Science 302 (5652), 1960-1963 (2003)
PUBMED   14671302
REFERENCE 2 (bases 1 to 1436)
AUTHORS  Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferrlera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
TITLE    Direct Submission
JOURNAL  Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT  This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES
source   1..1436
         /organism="Mus musculus"

```

```

/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1436
/gene="CHRNA3"
/locus_tag="HCML369"

ORIGIN

Alignment Scores:
Pred. No.: 1,336-73 Length: 1436
Score: 913.50 Matches: 187
Percent Similarity: 59.19% Conservative: 90
Best Local Similarity: 39.96% Mismatches: 172
Query Match: 34.28% Indels: 19
DB: 29 Gaps: 7

US-09-303-232-4 (1-496) x AY402875 (1-1436)
QY      34 GLuLyArGleLeuLeuHisLeuLeuAspHisTyrAsnValLeuGluArgProValVal 53
Db      21 GAGCAGCGCTGTTCAGTACCTGTTTGAAGATTACACGAGATCATCCGCGGTGGCT 80
QY      54 AsnGluSerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleLeuAspVal 73
Db      81 AACGTGCTCCATCCTGTCTATCCAGTTTGAGGTGTCTATGCTCAGCTGGTGAAGTG 140
QY      74 AspGluLysAsnGlnLeuLeuThrAsnIleTrpLeuLysLeuGluTrpAsnAspMet 93
Db      141 GATGAAGTAAACAGATCATGTGAACCAACCTGTGGTGAAGCAAAATCTGGAATGATAC 200
QY      94 AsnLeuArgTrpAsnThrSerAspPheGlyGlyValIysAspLeuArgValProHis 113
Db      201 AAGCTGAATGGAAACCTCTGACTACCAAGGGGTGAGTTTCATCGAGTCCCTGCAGAG 260
QY      114 ArgLeuTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspSerThr 133
Db      261 AAGATCTGAAACACGACATCGTGTTCACACACGCGATGGGATTTCCAAAGTGAT 320
QY      134 TyrProThrAsnValValArgAsnAsnGlySerCysLeuTyrValProProGlyIle 153
Db      321 GACAAACCAAAAGCTCTACTCAAGTACACAGGAGAAAGTGCATTTGGATCCCTCGGCGATC 380
QY      154 PheLysSerThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCysGlu 173
Db      381 TTAAAGAGCTCAAGCAAAATCGATGACCTACTTCCGTTTGACTACCAAAACTGCACC 440
QY      174 MetLysPheGlySerTrpThrTyrAspGlyTyrGlnLeuAspLeuGlnLeuAspGlu 193
Db      441 ATGAAGTTCGGCTCTCTGCTCTACGACAAAGCAAGATCGACCTGGTCTCTCATTTGGC-- 497
QY      194 GlyGlyAspLysSerSerPheValThrAsnGlyGluTrpGluLeuIleGlyValPro 213
Db      498 TCCTCCATGAACCTCAAGGACTATTGGGAAAGTGGCGAGTGGGCCATCATTAAGCCCCG 557
QY      214 GlyLysArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThrPhe 233
Db      558 GGTACAAACATGAATCAAGTACAACTCTCTGTGAGGAGATCTACCAAGACATCAGTAC 617
QY      234 AlaValIleArgArgLysThrLeuTyrTyrPhePheAsnLeuIleValProCysVal 253
Db      618 TCGCTATACATTCCGCGCTCGCTGTCTTACACCATCAACCTCATTCCTCGGTGCTG 677
QY      254 LeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeu 273
Db      678 CTCTATCTCTCTCTCTGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 737
QY      274 SerLeuGlyValThrIleLeuSerLeuThrValPheLeuAsnMetValAlaGluThr 293
Db      738 ACCTCTGATCTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 797
QY      294 MetProAlaThrSerAspAlaValProLeuLeuGlyThrTyrPheAsnCysIleMetPhe 313
Db      798 ATCCCTTCCACCTCACTGGTCTATCCCTTGTATCGGGGAGTACCTCTCTCTCTCTCTCT 857

```



```
Qy 312 MetPheMetValAlaSerSerValValSerThrIleLeuIleLeuAsnTyrHisHisArg 331
Db 852 ATGTTGTCGTGCACCTTTTCATCGTCACCACTGTGTGTCTCCTCAATGTGCACACCGT 911
Qy 332 HisAlaAspThrHisGluMetSerAspTyrIleArgCysValPheLeuTyrTrpLeuPro 351
Db 912 TCACCCAGCACTCACACATGGCATCTCTGGTCAAGGAGTGTCTTCCTGCACCACTGCC 971
Qy 352 TrpValLeuArgMetSerArgProGlySerAlaThrThrProProAlaAlaArgValPro 371
Db 972 ACCTTCCTCTTCATGAAGCGCCTGGTCTGGAAGTTCAGC-----CCGCGCAGGGTCCCT 1025
Qy 372 ProProProAspLeuGluLeu-----ArgGluArgSerSerIysSerLeuLeu----- 387
Db 1026 CATTCCAGCCAGCTGCACCTTGACACAGCTGAAGCTACGTCACCTCTGCGCTTAGGGCCCC 1085
Qy 388 -----AlaAsnValLeuAspIleAspAspPheArgHisProGlnAlaGlnGln 404
Db 1086 AGCAGCCCATCCACCTCTATGGAAATTCATGTACTTTGTGAACCTGTCCCTGCCACT 1145
Qy 405 ProGln---CysCysArgTyrTyrArgGlyGlyGluGluAsnGlyAlaGlyLeuAlaAla 423
Db 1146 CTTAAGTCTGAGTCAGTCAGCTCCACACGCGAGGCTCCCGAGGATGCCCGGCTGAGGTCC 1205
Qy 424 HisSerCysPheGlyValAspTyrGluLeuSerLeuIleLeuLysGluIleArgValIle 443
Db 1206 TCTGGGAGGTTCCGGCAAGAT-----CTACAGGAAGCATTAGAGGCGCTCAGCTTCATC 1259
Qy 444 ThrAspGlnMetArgLysAspAspGluAspAlaAspIleSerArgAspTyrLysPheAla 463
Db 1260 GCACAGCATTTGGAGATGATGTCGAGATCAGAGTCATCGAGGACTGGAAATTCGTC 1319
Qy 464 AlaMetValValAspArgLeuCysLeuIleIlePheThrLeuPheThrIleAlaThr 483
Db 1320 GCAATGTCGTGCGACCGCTGTTCTCTGGGTGTCTGTATGTGTGTTCTTCTGGGCACC 1379
Qy 484 LeuAlaValLeuLeu 488
Db 1380 ATGGGGCTCTTCCTG 1394

RESULT 12
AY402876
LOCUS
DEFINITION Homo sapiens CHRN4 gene, VIRUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION
VERSION AY402876
KEYWORDS GSS.
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1442)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civeillo,D.R., Lu,F., Murphy,B.,
Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
JOURNAL
PUBMED 14671302
REFERENCE
2 (bases 1 to 1442)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civeillo,D.R., Lu,F., Murphy,B.,
Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
Location/Qualifiers
```

```
source 1. .1442
/mol_type="taxon:9606"
/db_xref="taxon:9606"
<1. ->1442
/gene="CHRN4"
/locus_tag="HCM1370"

ORIGIN
Alignment Scores:
Pred. No.: 6.24e-72 Length: 1442
Score: 895.50 Matches: 184
Percent Similarity: 58.60% Conservative: 92
Best Local Similarity: 39.07% Mismatches: 163
Query Match: 33.60% Indels: 32
DB: 29 Gaps: 7

US-09-303-232-4 (1-496) x AY402876 (1-1442)
Qy 34 GluLysArgLeuLeuHisHisLeuLeuAsp-----HisTyrAsnValLeuGluArgPro 51
Db 24 GAGGAAAAGCTGATGGACGACCTTCTGACAAAAACCGTTACATAAACCTGATCGGCCA 83
Qy 52 ValValAsnGluSerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleIle 71
Db 84 GCCACCAAGCTCTCACAGCTCATCTCCATCAAGCTGCAGCTCTCCCTGGGCCAGCTTATC 143
Qy 72 AspValAspGluLysAsnGlnLeuLeuIleThrAsnIleTrpLeuLysLeuGluTrpAsn 91
Db 144 AGCGTGAATGAGCGAGCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 203
Qy 92 AspMetAsnLeuArgTrpAsnThrSerAspPheGlyGlyValLysAspLeuArgValPro 111
Db 204 GATTACCGCGCTGATCTGGAACAGCTCCCGCTACGAGGGTGTGAACATCTCTGAGATCCCT 263
Qy 112 ProHisArgLeuTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAsp 131
Db 264 GCAGAGCGCATCTGGTTCGCTGATCGCTGCTTTACAAACGCCGAGCGGACCTATGAG 323
Qy 132 SerThrTyrProThrAsnValValArgAsnAsnGlySerCysLeuTyrValProPro 151
Db 324 GTGCTCTGTACACCACTTGTAGTCCGCTCCAAACGCGAGCGTCTGTGGTCCCTCCCT 383
Qy 152 GlyIlePheLysSerThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArg 171
Db 384 GCCATCTACAAGAGCGCTGCAAGATTGAGTGAAGTACTTCTCCCTTCGACGACGAGAAC 443
Qy 172 CysGluMetLysPheGlySerTrpThrTyrAspGlyTyrGlnLeuAspLeuGlnLeuGln 191
Db 444 TGCACCTCAAGTTCGCTCTCTGACCTATGACACACGAGGATGAGATGGTCTCATG 503
Qy 192 AspGluGlyGlyAspIleSerSerPheValThrAsnGlyGluTrpGluLeuIleGly 211
Db 504 ACGCCC--ACAGCCAGCATGGATGACTTTTACTCCCGAGTGTGTGAGTGGGACATAGTGCC 560
Qy 212 ValProGlyLysArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIle 231
Db 561 CTCCAGGGAGAGGACA-----GTGAACCCCAAGACCCCGACCTACGTGGAGCGTG 611
Qy 232 ThrPheAlaValValIleArgArgLysThrLeuTyrTyrPhePheAsnLeuIleValPro 251
Db 612 ACTTACGACTTCATCATCAAGCGCAGAGCTCTGTCTTACACCATCAACCTCATCATCCCC 671
Qy 252 CysValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProAspSerGlyGlu 271
Db 672 TGGTGTCTACCACTTGTGGCCATCTCTTCTTACTCTGCATCCGACTGGGGGAG 731
Qy 272 LysLeuSerLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAla 291
Db 732 AAGATGACATGTGCATCTCAGTGTCTGCTGGCACTGACATTTCTTCTGTCTCATCTCC 791
Qy 292 GluThrMetProAlaThrSerAspAlaValProLeuLeuGlyThrTyrPheAsnCysIle 311
Db 311
```


Db 1291 CTGGAGTCTTTATGCTGGTGTCTCATCGGACGCTGGCTGTG 1335

Search completed: May 9, 2004, 04:20:29
Job time : 3493.3 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: May 7, 2004, 12:59:49 ; Search time 5953.59 Seconds
(without alignments)
10942.062 Million cell updates/sec

Title: US-09-303-232-5_COPY_95_1597
Perfect score: 1503
Sequence: 1 atggccctatgttggcggc.....caccgcataatcatcgtgcaa 1503

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

- GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*
29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htgo.hum.*
40: em.htgo.mus.*
41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1503	100.0	3029	3	AF143847	AF143847 Heliothis
2	1503	100.0	3109	6	E58348	E58348 Nucleic aci
3	1503	100.0	3109	6	AX009614	AX009614 Sequence
4	770.2	51.2	2023	3	AF321447	AF321447 Drosophil
5	751	50.0	2023	3	AF321445	AF321445 Drosophil
6	751	50.0	2023	3	AF321446	AF321446 Drosophil
7	747.8	49.8	1865	3	DME554209	AJ554209 Drosophil
8	747.8	49.8	1865	3	BT011147	BT011147 Drosophil
9	715.2	47.6	2068	3	AF321448	AF321448 Drosophil
10	654	43.5	2110	3	AF321449	AF321449 Drosophil
11	575	38.3	3629	3	AF143846	AF143846 Heliothis
12	575	38.3	3701	6	AX009612	AX009612 Sequence
13	575	38.3	3701	6	E58347	E58347 Nucleic aci
14	535	35.6	1683	3	DME554210	AJ554210 Drosophil
15	521.6	34.7	2307	3	AF272778	AF272778 Drosophil
16	412.4	27.4	2834	3	AY036613	AY036613 Drosophil
17	411.4	27.4	1668	5	AY298752	AY298752 Takifugu
18	411.4	27.4	2886	6	E58346	E58346 Nucleic aci
19	411.4	27.4	2886	6	AX009610	AX009610 Sequence
20	378.8	25.2	2107	9	AF486623	AF486623 Macaca mu
21	377.2	25.1	1509	9	HSNACHRA7	Y08420 H.sapiens m
22	377.2	25.1	1916	9	AF385585	AF385585 Homo sapi
23	377.2	25.1	1977	9	HSU40583	U40583 Human alpha
24	376.8	25.1	1551	4	BTAT7	X93604 B.taurus mr
25	376.4	25.0	2106	10	RATNAPAD	L31619 Rattus ratt
26	376.4	25.0	3030	10	HSARA7A	S53987 nicotinic r
27	375.6	25.0	2087	9	HUNA7NAR	X70297 H.sapiens m
28	375	25.0	1559	9	AX054567	L25827 Human a7 ni
29	374	24.9	1509	6	AX054567	AX054567 Sequence
30	374	24.9	1876	6	AR282833	AR282833 Sequence
31	374	24.9	1876	6	AX719088	AX719088 Sequence
32	374	24.9	1876	9	HSU62436	U62436 Human nicot
33	374	24.9	2088	10	AF225980	AF225980 Mus muscu
34	373.6	24.9	1590	6	AR261850	AR261850 Sequence
35	373.6	24.9	1590	6	BD023656	BD023656 Variant h
36	373.6	24.9	1876	6	AR055255	AR055255 Sequence
37	373.6	24.9	1876	6	AR071403	AR071403 Sequence
38	373.6	24.9	1876	6	AR173187	AR173187 Sequence
39	373.6	24.9	1876	6	AR224030	AR224030 Sequence
40	372.4	24.8	1509	6	AX054575	AX054575 Sequence
41	372.4	24.8	1509	6	AX054577	AX054577 Sequence
42	372.4	24.8	1848	10	MUSNARS	L37663 Mus musculu
43	371.8	24.7	1665	5	AY298753	AY298753 Takifugu
44	370.8	24.7	1509	6	AX054579	AX054579 Sequence
45	365.8	24.3	1542	5	AY299466	AY299466 Takifugu

ALIGNMENTS

RESULT 1
AF143847
LOCUS AF143847
DEFINITION Heliothis virescens putative nicotinic acetylcholine receptor alpha
7-2 subunit mRNA, complete cds.
ACCESSION AF143847
VERSION AF143847.1 GI:4895006
KEYWORDS
SOURCE Heliothis virescens (tobacco budworm)
ORGANISM Heliothis virescens
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuoidea; Noctuidae; Heliothinae; Heliothis.
REFERENCE 1 (bases 1 to 3029)
AUTHORS Schulte,T., Oellers,N. and Adamczewski,M.


```
VERSION AX009614.1 GI:9996846
KEYWORDS
SOURCE Heliothis virescens (tobacco budworm)
ORGANISM Heliothis virescens
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuoidea; Noctuidae; Heliothinae; Heliothis.
REFERENCE 1
AUTHORS Adamczewski, M.D., Schulte, T.D. and Oellers, N.D.
TITLE Nucleic acids encoding acetylcholin-receptor subunits from insects
JOURNAL Patent: EP 0962528-A 5 08-DEC-1999;
BAYER AG (DE)
FEATURES
    source
        1..3109
            /organism="Heliothis virescens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:7102"
            /note="unnamed protein product"
            /codon_start=1
            /protein_id="CAC07501.1"
            /db_xref="GI:9996847"
            /db_xref="REMTREMBL:CAC07501"
            /translation="MAEMLAALALLPVSQBPHEKRLNALLNANTLERPVANE
            SELEVFRLTQIIDVDEKNQLLIINWLEWNYNLRWNSDEYGVKDLRIITPN
            KLMKPDVLYNSADGEDGTQYNVVRSGSLVPPGIFKSKCKMDIAWFPDDOH
            CDMKFGSYTDNQLDLVXDEAGDLSDRITNGEWFLIGMPGKNITITVACCPVPV
            DVNFTIMRERTLYFENLJVPCLVLISSMALLGFTLPDPSGKLTGLVTLISLTVFL
            NLVATILPQVSDAIPLLGTFCNFMFVASSVLTVVLYNHHRTADIHENPOWIKSV
            FLOWPLMIRSPGKIKTRKIMMTRMRELEKRSKSLANLVIDDDFRHGPP
            PPNSTAGNLPGCSIFTRFRRSPVPMEDVGGGLSHHREHLILRELFQITA
            RMKKADBEALISDWKFAAMVVDVRFCLFVETLFTIATVAVLLSAPHIIQV"
    CDS
        100.0%; Score 1503; DB 6; Length 3109;
        Best Local Similarity 100.0%; Pred. No. 0;
        Matches 1503; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    ORIGIN
        1 ATGCCCCCTATGTTGGCGGCTTGGCGCTGTGCTTGTGCTGCCGCTATCGGACCAAGT 60
        95 ATGCCCCCTATGTTGGCGGCTTGGCGCTGTGCTTGTGCTGCCGCTATCGGACCAAGT 154
        61 CCTCAGAGAGAGACTCTGAACGCTTGTGCGGAACACTACAACCCCTGGAGCGACCG 120
        155 CCTCAGAGAGAGACTCTGAACGCTTGTGCGGAACACTACAACCCCTGGAGCGACCG 214
        121 GTGCGCAACGAGAGCGGACCGCTAGAGTCAGGTTCCGCTTGCCTTGCAGCAATCAT 180
        215 GTGCGCAACGAGAGCGGACCGCTAGAGTCAGGTTCCGCTTGCAGCAATCAT 274
        181 GAGCTGACGAGAGAGAACTCACTACTTATAACCAATATATGGCTGTGCTTGGAGTGAAT 240
        275 GAGCTGACGAGAGAGAACTCACTACTTATAACCAATATATGGCTGTGCTTGGAGTGAAT 334
        241 GACTACAACTGAGTGGAAACGACGAGATATGGCGGGTCAAGGACCTCAGGATCAGC 300
        335 GACTACAACTGAGTGGAAACGACGAGATATGGCGGGTCAAGGACCTCAGGATCAGC 394
        301 CCCAAACAAGTTGTGGAAGCGGACGCTCTTATGTAATAGTGTGACGAGGGTTTGAC 360
        395 CCCAAACAAGTTGTGGAAGCGGACGCTCTTATGTAATAGTGTGACGAGGGTTTGAC 454
        361 GGGACCTTACAGACCAACGTTGGTTCAGAAAGCGGGCGGAGTTGCCCTGTACGTGCCACCT 420
        455 GGGACCTTACAGACCAACGTTGGTTCAGAAAGCGGGCGGAGTTGCCCTGTACGTGCCACCT 514
        421 GGCATATTCAAGAGCACATCGAGATGGACATCGCGTGGTTCCCTTCGACGACCAACAC 480
        515 GGCATATTCAAGAGCACATCGAGATGGACATCGCGTGGTTCCCTTCGACGACCAACAC 574
        481 TGTGATATGAAGTTTCGCTAGCTGGACATATGACGGCAATCAGTTGGATCTGGTCTAAAA 540
        575 TGTGATATGAAGTTTCGCTAGCTGGACATATGACGGCAATCAGTTGGATCTGGTCTAAAA 634
```

```
QY 541 GATGAGCAGCGCGGATCTATCGGACTTCTATACAAATCGGAGTGGTATCTAATAGA 600
DB 635 GATGAGCAGCGCGGATCTATCGGACTTCTATACAAATCGGAGTGGTATCTAATAGA 694
QY 601 ATGCCAGGCAAAAAGAACACAATAACATACGCGTGTGCGCCGAGCCCTACGTGACGTC 660
DB 695 ATGCCAGGCAAAAAGAACACAATAACATACGCGTGTGCGCCGAGCCCTACGTGACGTC 754
QY 661 ACCTTCACCATCATGATAAGAACGAACTTGTGTACTTCTTCAACTGATCGTCCG 720
DB 755 ACCTTCACCATCATGATAAGAACGAACTTGTGTACTTCTTCAACTGATCGTCCG 814
QY 721 TGGTGTGTGATCTCATCGATGGCACTCTCGGCTTCACTGCGCACAGATCCGGAGAG 780
DB 815 TGGTGTGTGATCTCATCGATGGCACTCTCGGCTTCACTGCGCACAGATCCGGAGAG 874
QY 781 AAACCTCACACTTGGAGTCACATATTCTTCTATCGCTGACGGTGTTCCTCAACTGATGCC 840
DB 875 AAACCTCACACTTGGAGTCACATATTCTTCTATCGCTGACGGTGTTCCTCAACTGATGCC 934
QY 841 GAGACCTGCCACAGGTCCTCGAGCTATCCCTCTGTAGGACGTACTTCAATTGCAATC 900
DB 935 GAGACCTGCCACAGGTCCTCGAGCTATCCCTCTGTAGGACGTACTTCAATTGCAATC 994
QY 901 ATGTTCAATGTTAGCGTCTGTGCTACTGACTGTGCTGTGCTACTCAATTACCAATCGA 960
DB 995 ATGTTCAATGTTAGCGTCTGTGCTACTGACTGTGCTGTGCTACTCAATTACCAATCGA 1054
QY 961 ACAGCTGATATACATGAAATGCCACAGTGGATAAAATCAGTATTCTTACAAATGGTTGCCA 1020
DB 1055 ACAGCTGATATACATGAAATGCCACAGTGGATAAAATCAGTATTCTTACAAATGGTTGCCA 1114
QY 1021 TGGATCTCGAATGTGAGCGCCAGGAAAGAAATCACCGAAGAACTATAATGATGAAC 1080
DB 1115 TGGATCTCGAATGTGAGCGCCAGGAAAGAAATCACCGAAGAACTATAATGATGAAC 1174
QY 1081 ACAGGAGATGAGGAGCTGGAACCTGAAGGAGAGGTCGTGGAAGTCTTGTGCGCAAGT 1140
DB 1175 ACAGGAGATGAGGAGCTGGAACCTGAAGGAGAGGTCGTGGAAGTCTTGTGCGCAAGT 1234
QY 1141 CTAGATATTGATGATGACTTTCAGACACGGCCCTCCGCTCTTAACAGTACTGCTCGAC 1200
DB 1235 CTAGATATTGATGATGACTTTCAGACACGGCCCTCCGCTCTTAACAGTACTGCTCGAC 1294
QY 1201 GGGAAATTTGGGACCTGGGTGCTCAATATTCCGACCGAATTTCCGTCGGTCTGCTCGCT 1260
DB 1295 GGGAAATTTGGGACCTGGGTGCTCAATATTCCGACCGAATTTCCGTCGGTCTGCTCGCT 1354
QY 1261 CCCTCCACGATGGAAGACGTGGCGGGCTGGGTAGCCACCATCGGAGCTGCACCTC 1320
DB 1355 CCCTCCACGATGGAAGACGTGGCGGGCTGGGTAGCCACCATCGGAGCTGCACCTC 1414
QY 1321 ATACTGAGAGAGCTGCAGTTTCATCAGCGCCAGGATGAAGAGGCTGATAGGAGCCGAG 1380
DB 1415 ATACTGAGAGAGCTGCAGTTTCATCAGCGCCAGGATGAAGAGGCTGATAGGAGCCGAG 1474
QY 1381 CTGATCAGGACTGGAAGTTTCTGCGATGGTTTGTGATAGGTTTGGCTGTGCTGCTG 1440
DB 1475 CTGATCAGGACTGGAAGTTTCTGCGATGGTTTGTGATAGGTTTGGCTGTGCTGCTG 1534
QY 1441 ACATTTTTCACATCATCGGACAGTAGCTGTCTGTTATCGGCACCGCATATCATCTG 1500
DB 1535 ACATTTTTCACATCATCGGACAGTAGCTGTCTGTTATCGGCACCGCATATCATCTG 1594
QY 1501 CAA 1503
DB 1595 CAA 1597
```

RESULT 4
AF321447
LOCUS

2023 bp mRNA linear INV 29-APR-2002

DEFINITION Drosophila melanogaster nicotinic acetylcholine receptor Dalpha6 subunit variant type III (nAcRalpha-30D) mRNA, complete cds, alternatively spliced.

ACCESSION AF321447

VERSION AF321447.1 GI:20152848

KEYWORDS

SOURCE

ORGANISM Drosophila melanogaster (fruit fly)

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE 1 (bases 1 to 2023)

AUTHORS Grauso, M., Reenan, R.A., Culetto, E. and Sattelle, D.B.

TITLE Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes, Dalpha5, Dalpha6 and Dalpha7, in Drosophila melanogaster Identify a New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-Mediated A-to-I Pre-mRNA Editing

JOURNAL Genetics 160 (4), 1519-1533 (2002)

MEDLINE 21969411

PUBMED 11973307

REFERENCE 2 (bases 1 to 2023)

AUTHORS Grauso, M. and Sattelle, D.B.

TITLE Direct Submission

JOURNAL Submitted (15-NOV-2000) Human Anatomy and Genetics-Functional Genetics Unit, University of Oxford, South Park Road, Oxford OX1 3QX, UK

FEATURES

source

Location/Qualifiers

1..2023

/organism="Drosophila melanogaster"

/mol_type="mRNA"

/db_xref="taxon:7227"

/chromosome="2"

/map="30D1"

/dev_stage="embryo"

1..2023

/gene="nAcRalpha-30D"

379..1863

/gene="nAcRalpha-30D"

/note="ion channel; neurotransmitter transmembrane receptor; alternatively spliced; contains exons 3b and 8b"

/codon_start=1

/product="nicotinic acetylcholine receptor Dalpha6 subunit variant type III"

/protein_id="AA013394.1"

/db_xref="GI:20152849"

/translation="MDSFLPASLSFLVLLFLALIKESCOGPHKRLNLLSTNYTL ERVANESEPLEVKFGTLQOIIDVDEKNQLLTNLSLEWINDYLRWNETYGVKDLRITPNKLPVLYMNSADBGFDGTHTNIVVXHSGLYVPPGIFKSKCRMDITWFPFDQHCENKMGFSWTYDGNLDLVNSEDGDLSDFITNGEWYLLAMPCKNTIVVACPEPYVDIITFTQIRRTLYYFNLIVPCVLISSMALLGFTLPPDSGEKLTGLVTILLSLTVELNLVAETLPQVSDAIPILGYFCINMFWASSVLTIVVNLNHHRTADIHHEMPPMKSVFLQWLFWLWRGPRKRTIKRTILLNRMKELEKRSKSLLANVLDIDD DFRHTISGQTALGSSASFORPTTVEHHTAIGCNKDLHLILKELOFTARMRKADD EAEILIGDWKFAAMVDRFLIVFTLFTIATVTLVSAPHIIQV"

794

/gene="nAcRalpha-30D"

/note="results in glycine to serine substitution; compared to the sequence deposited in GenBank Accession Number AE003626"

846

/replace="a"

/gene="nAcRalpha-30D"

/note="results in asparagine to serine substitution; compared to the sequence deposited in GenBank Accession Number AE003626"

846

/replace="a"

variation

variation

ORIGIN

Query Match 51.28; Score 770.2; DB 3; Length 2023;

Best Local Similarity 70.94; Pred. No. 1.1e-193;

Matches 1058; Conservative 0; Mismatches 393; Indels 42; Gaps 1;

QY 11 TGTGGCGGCTTGGCGCTTGTGCTTGTGCTGCGCGCTTCGGAGCAAGGTCTCTAGATTG 1150

Db 410 TGTGTGCTCTGTTGATCTTTCTGCGATAATTAAGAAAGCTGTCAAGGACCTCATGAAA 469

QY 71 AGAGACTCTCTGAACGCGTGTCTGCGAAGTAAACACACCTCGAGCGACCGGTGGCAACG 130

Db 470 AGCGCTGCTGAACCACTCTGCTGCCCTACATACGCTGAGGAGCCCGTGGCAATG 529

QY 131 AGACGGAACCGCTAGAGTTCAGGTTCGGCTTGACCTTGCAGCAATATTCACGTGGAG 190

Db 530 AATCGGAGCCCGCTGGAGGTTAAGTTCGACCTGACGCTGCAGCAGATCATCGACGTGAG 589

QY 191 AGAAGATCAACTACTTATTAACCAATATATGCTGTGCTGTGGAGTGGATCACTACACC 250

Db 590 AGAAGATCACTCTCTCATACGAATCTTTGGCTTTTGGTGGAGTGGAGCACTACATC 649

QY 251 TGAGGTGGAACGACAGCGAGTATGCGGGGTCAAGGACCTCAGGATCAGCCCAACAGT 310

Db 650 TGGCTGGAATGAACGAATACGCGGGGTCAAGGATCTACGAATCAGGCCCAACAGC 709

QY 311 TGTGGAACCGGACGCTCTTATGTATATAGTGTGACGAGGGTCTTTCACGGGACCTACC 370

Db 710 TGTGGAACCGGACGCTCTCATGTACAACAGCGCGGATGAGGATTCGATGCCAGTATC 769

QY 371 AGACCAACGCTGCTCAGHAGCGCGGCGAGTTCCTCTGCTGCTGCTGCTGCTGCTGCT 430

Db 770 ACACCAACATTTGCTGCTCAACATAGCGGCGAGTGTCTGTAGCTGCGCCCTGCTATCT 829

QY 431 AGACCAACATGAGATGAGATCGCGTGTGCTTCCCTTCGACGACCAACACTGTCTATGA 490

Db 830 AGACCAATGCAAGATGAGATCAGCTGTGCTTCCCTTCGACCAATTCGAAATGA 889

QY 491 AGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 550

Db 890 AATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 949

QY 551 GCGCGCTCTATCGGCTTCTATCAACATGCGGAGTGGAGTGGATCTATAGGATGCGAGCA 610

Db 950 GAGGGGATCTTTCGATTTTCAACAAATGCGGAGTGGATCTTCTGCTTGGCTTGGCTAA 1009

QY 611 AAAAGAACACAAATAACATAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 670

Db 1010 AGAAGATACGATAGTCTAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1069

QY 671 TCATGATAGAGAGCAACCTTGTACTACTTCTTCAACCTGATGCTGCTGCTGCTGCTGCT 730

Db 1070 TACAAATTCGTCGCGTACATATATATATATATATATATATATATATATATATATAT 1129

QY 731 TCTCATGATGCGCACTCTCGGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 790

Db 1130 TCTCATGATGCGCACTCTCGGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1189

QY 791 TTGAGTCACTATTCTTCTATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 850

Db 1190 TGGAGTTACAAATCTTCTATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1249

QY 851 CAGAGTCTCGGAGTATCCCTCTGTTAGGAGTATCTTCAATTCATGCTGCTGCTGCTGCT 910

Db 1250 CCCAAGTATCGATGCAATCCCTCTGTTAGGAGTATCTTCAATTCATGCTGCTGCTGCT 1309

QY 911 TAGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 970

Db 1310 TCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1369

QY 971 TACATGAAATGCCACAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1030

Db 1370 TTCAAGGATGCCACCGTGGATCAAGTCCGTTTCTTCAATGCTGCTGCTGCTGCTGCTG 1429

QY 1031 GAATGTCGAGGCGGAGGAGAGATCAACAGGAGGATCAATGATGATGATGATGATGATGAT 1090

Db 1430 GAATGCTGAGCGCGTGGCAAGATTTACAGCAAAACAACTATATTAAGCAATCGCATGA 1489

QY 1091 GCGAGCTGGAACCTGAGGAGAGGCTGCTGAGTCTTCTGCGGAGTGTCTTCTAGATTG 1150


```
QY 371 AGACCAACGTGGTGTGAGAGCGGGCGAGTTCCTGTAGCTGCGACCTGGCATATTC 430
Db 770 ACACCAACATTTGGTGTGAAACATACCGCGAGTGTCTGTACGTGCCCCCTGGTATCTCA 829
QY 431 AGAGCACATGCAAGATGAGATCGCGTGGTTCCTCTCGACGACCAACACTGTGATGA 490
Db 830 AGAGCACATGCAAGATGAGATCAGCGTGGTTCCTCATTTGATGACCAACATTCGGAATGA 889
QY 491 AGTTCCGTAGTGGACATATGACGGCAATCAGTTGGATCTGGTGTCTAAAGATGAGCGAG 550
Db 890 AATTCCGTAGTGGACTTACGATGGAATCAGTTGGATTTGGTTTGAATTCGGAAGATG 949
QY 551 GCGCGCATCTATCGGACTTCATACAAATGGGAGTGTATCTAAATAGGATGCGAGGCA 610
Db 950 GAGGGATCTTTCCGATTTTATACAAATGGGAGTGTACTTGTCTTGCATGCCGCGGAA 1009
QY 611 AAAGAAACAAATAACATACCGGTGTCGCCCGAGCCCTACGTGGAGCTCACCTTCACCA 670
Db 1010 AGAAGATACGATAGTCTACGCTCTGCTGCCAGAACCATATGTGATATCACCTTTACTA 1069
QY 671 TCATGATAAGAGACGACCTTGTACTACTTCTCAACCTGATGCTCCGCTGGTGTGA 730
Db 1070 TACAAATTCGTGCGCGTACATATATATTTTCAATTTTAATCGTCCATGTGTGCTAA 1129
QY 731 TCTCATCGATGGCACTCTCGGCTTTCACACTGCCACACAGACTCCGGAGAGAACTCACAC 790
Db 1130 TCTCATCGATGGCCCTACTCGGCTTTCACATGCGCGCGGATTCGGCGAGAACTGACGC 1189
QY 791 TTGGAGTCACTATTTCTTCTATCGCTGACGGTGTTCCTCAACCTGGTAGCCGAGACCTG 850
Db 1190 TGGGAGTTTCAATTTCTTCTATCGCTACAGTGTTCCTCAACCTTGTAGCTGAGACATTC 1249
QY 851 CACAGTCTCCGAGCTATCCCTCTTAGGAGGTACTTCAATGTGATCATGTTATGG 910
Db 1250 CCAAGTATCTGATGCAATCCCTCTTAGGAGCTTCAATGTGATCATGTTATGG 1309
QY 911 TAGCGTCTGTGGTACTGACTGTGGTGTGTACTCAATTAACCACTCGAAGCAAGCTGATA 970
Db 1310 TCGCTCTGTGGTGTGCTGACAGTAGTGTGTCTCACTACCACTCGCAGCGGAC 1369
QY 971 TACATGAATGCCACAGTGGATAAATCAGTATTTCTTACAAATGGTTGCCATGGATCTGC 1030
Db 1370 TTCAGAGATGCCACGTGGATCAAGTCCGTTTCTTCAATGGCTGCTCCCTGGATCTGC 1429
QY 1031 GAATGTCGAGCCAGGAGAAAGATCACAGGAGAGACTTAATGATGAACAGAGATGA 1090
Db 1430 GAATGGTTCGACCGGTGCGAAGATTACAGCAAAACAATACTATTAAAGCAATCGCATGA 1489
QY 1091 GCGAGCTGGAACCTGAAGGAGAGGTGCTGGAAGTCTTGTGCGGAATGTTCTAGATATTG 1150
Db 1490 AGGAGCTGGAGCTTAAGGAGCGCTCTCCAAATCCCTGTGCGCAATGTCTCGACATCG 1549
QY 1151 ATGATGACTTCAGACAGCGGCCCTCGGCTCTTAACAGTACTGCTCGACCGGAAATTGG 1210
Db 1550 ACAGACACTTCCGCGCACAAATATCTG----- 1576
QY 1211 GACCTGGTGTCTAAATATTCGACAGGATTTCCGTGGTGGTTCGTCCGTCCGTCACGA 1270
Db 1577 -----GCTCCCAACCGCATTTGGCTGCTGCGGCAAGCTTCGTGCGGCCACCAACGG 1627
QY 1271 TGAAGACGTGGGCGCGGCTGGGTAGCCACCATCGCAGCTGCACCTCATACTGAGAG 1330
Db 1628 TGGAGGAGCATCACACGGCCATCGGCTGGCAATCACAAGATCTTCACTAATTTCTCAAG 1687
QY 1331 AGTGCAGTTTCATACGCGCAGGATGAAGAGGCTGATGAGGAGCGGAGCTGATCAGCG 1390
Db 1688 AATTGCAATTTATTACGGCGCGATGCGCAAGCTGACGACGAGCGGAATTCATCGCG 1747
QY 1391 ACTGGAAGTTTGTGCGATGGTTGTGATAGGTTTTCCTGTTGCTGTTTCACTTTTCA 1450
Db 1748 AATTGGAAGTTTCGCGGCAATGGTTGTGATAGATTTTGTAAATTTGTTTTCACGCTTCA 1807
QY 1451 CAATCATCGGACAGTAGTGTCTCTGTTATCGGACCGCATATCATGTCGCA 1503
```

```
Db 1808 CGATTATTGCAACGGTTACGGTGTCTCTCCGCTCCGCACATAATCGTGCAA 1860
RESULT 6
AF321446 2023 bp mRNA linear INV 29-APR-2002
LOCUS Drosophila melanogaster nicotinic acetylcholine receptor Dalpha6
DEFINITION subunit variant type II (nAcRalpha-30D) mRNA, complete cds,
alternatively spliced.
ACCESSION AF321446
VERSION AF321446.1 GI:20152846
KEYWORDS Drosophila melanogaster (fruit fly)
SOURCE Drosophila melanogaster
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 2023)
AUTHORS Grauso,M., Reenan,R.A., Culetto,E. and Sattelle,D.B.
TITLE Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
Dalpha5, Dalpha6 and Dalpha7, in Drosophila melanogaster Identify a
New and Highly Conserved Target of Adenosine Deaminase Acting on
RNA-Mediated A-to-I Pre-mRNA Editing
JOURNAL Genetics 160 (4), 1519-1533 (2002)
MEDLINE 21969411
PUBMED 11973307
REFERENCE 2 (bases 1 to 2023)
AUTHORS Grauso,M. and Sattelle,D.B.
TITLE Direct Submission
JOURNAL Submitted (15-NOV-2000) Human Anatomy and Genetics-Functional
Genetics Unit, University of Oxford, South Park Road, Oxford OX1
30X, UK
FEATURES
source
1..2023
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/chromosome="2"
/map="30D1"
/dev_stage="embryo"
1..2023
/gene="nAcRalpha-30D"
379..1863
/gene="nAcRalpha-30D"
/notes="ion channel; neurotransmitter transmembrane
receptor; alternatively spliced; contains exons 3b and 8a"
/codon_start=1
/product="nicotinic acetylcholine receptor Dalpha6 subunit
variant type II"
/protein_id="AAM13393.1"
/db_xref="GI:20152847"
/translation="MDSPLPASLSLFLVLLIFLAIKESCGPHEKLLNHLSTYNTL
ERPVANESPLEVKFGLTLQIIVDDEKNQLITNLWLSLEWYDNLRNWETYGVK
DLRITPNKLPKDLVLYNSADSGDGTGHTNLVVKHSGSLVVPVPGI EKSTCKMDITW
FPFDQHCMEKGSWTYDGNQLDLVLSNDSGDLSDFTNGEWYLLAMPCKNTIYVA
CPEKVDIIFITQIRRTTYFFNLIVPCVLSSWALIGFTLPPDSGSKLTUGVTIL
LSLTVELNLVAESMPTSDAPLIGTYFCIMFMVASSVLTAVVLYNHHRTADIHFM
PWIKSVLQWLPWILRMGRPKRTIKRTILLSNRMKELEKERSKSLANVLDD
DFRHTTSGSOTAGSSAFRPTTVEEHTAIGCNHKLHLILKELQFITARMRKADD
EAEELIGDKWFAAMVDRFCLIVFTLTITATVTLLSAPHIIIV"
793..794
/gene="nAcRalpha-30D"
/notes="results in glycine to asparagine substitution;
compared to the sequence deposited in GenBank Accession
Number AF003626"
/replace="aa"
794
variation
/gene="nAcRalpha-30D"
/notes="results in glycine to serine substitution; compared
to variant clone"
/replace="a"
846
variation
```

D	b	1130	TCTCATCGATGGCCCTACTGGGCTTCACATTGCCGGCGGATTTCGGGGAGAACTGACGC	1187
Q	y	791	TTGGAGTCACTATTCTTTCTATCGCTCAGCGTGTTCTCTCAACCTGGTAGCCGAGACCTGTC	850
D	b	1190	TGGGGGTAACTAFACTACTATCATTAACAAGTATTTCTAAACCCTGTGCGCGAGTCCATGC	1249
Q	y	851	CACAGGTCTCCGACGCTATCCCCCTGTTAGGGAAGTACTTCAATTGCATCATCTTCATGG	910
D	b	1250	CGAACACGTCGGATGCTGTTCTCTTATAGGCACTTCAATTGCATCATCTTCATGG	1309
Q	y	911	TAGCGTCTGTGTGTACTGACTGTGGTGATACTCAATTACCACTACGAACTGATG	970
D	b	1310	TCCGCTCGTGGTGTGTGACAGTAGTGTGTCTCACTACCACTACGAACTGATG	1369
Q	y	971	TACATGAATGCCACAGTGGATGATAAATCAGTATTCTCTCAATGGTGGCCATGATCTGC	1030
D	b	1370	TTCACGAGATGCCACGCTGGATCAAGTCCGTTTTTCTCAATGGCTGCCCTGGATCTTGC	1429
Q	y	1031	GAAATGTCGAGGCGCAGGAGAAGATCACAGGAAGACTATAATGATGAACACAGAGATGA	1090
D	b	1430	GAAATGGGTGCACCGGTCGCAAGATTACACGCAAAACAATCTATTAAAGCAATCGCATGA	1489
Q	y	1091	GGAGCTGGAACTGAAGGAGAGTGTGCGAAGTCTTGTGTGGCGAATGTTCTTAGATATTG	1150
D	b	1490	AGGAGCTGGAGCTAAAAGGAGCGCTCTCCAAAATCCCTGTGTGGCCAATGTCTCTGACATCG	1549
Q	y	1151	ATGATCACTTCACACACGCGCCCTCGCCCTCCTAACAGTACTGCTCGACCGGGAATTTGG	1210
D	b	1550	ACGACGACTTCCGGGCACAAATATCTG-----	1576
Q	y	1211	GACCTGGGTGCTCAAATATTCCGCACAGGATTTCCGTGGTTCGTGGTTCGGTCCGTCACAGA	1270
D	b	1577	-----GCTCCAAAACCGCCATTGGCTGCTGCGCAGCTTCGGTTCGGCCCACACACG	1627
Q	y	1271	TGGAAGACGTGGCGGGCTGGTAGCCACCATCGGAGCTGCACCTCATFACCTAGAG	1330
D	b	1628	TGGAGGAGCATCACACGGCCATCGGTGCAATCACAAGATCTTCAATTCTCAAAG	1687
Q	y	1331	AGCTGCAGTTTCATCAGCGCCAGCATGAAGAAGCTCATAGGAAGCGGAGCTCATCAGCG	1390
D	b	1688	AATTGCNAATTATATCGGCGCGATGCGCAAGCTGACGACGAAGCGGAATTGATCGCG	1747
Q	y	1391	ACTGGAAGTTGCTGCGATGGTTGTGATAGGTTTTTGCCTGTGTTTGCACACTTTTCA	1450
D	b	1748	ATTGGAAGTTTCGGCGGAATGGTTGTGGATAGATTTTGTTTAATTGTTTTTTCAGCTCTTCA	1807
Q	y	1451	CAATCATCGGACAGTAGCTGCTGTATTCGGCAGCGCATATCATCTGTGCAA	1503
D	b	1808	CGAATTATGCAAGTTACGGTGTCTGCTCTCCGCTCGGACATANTGTGCAA	1860
<hr/>				
RESULT 7				
DME554209			1665 bp mRNA linear INV 01-APR-2003	
LOCUS				
DEFINITION			Drosophila melanogaster mRNA for nicotinic acetylcholine receptor subunit Dalpha6 (nAcAlpha-30D gene).	
ACCESSION			AJ554209	
VERSION			AJ554209.1 GI:29466434	
KEYWORDS			nAcAlpha-30D gene; nicotinic acetylcholine receptor subunit Dalpha6.	
SOURCE			Drosophila melanogaster (fruit fly)	
ORGANISM			Drosophila melanogaster	
			Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	
REFERENCE			1 (bases 1 to 1665)	
AUTHORS			Millar,N.S.	
TITLE			Direct Submission	
JOURNAL			Submitted (31-MAR-2003) Millar N.S.; Department of Pharmacology,	
			University College London, Gower Street, London, WC1E 6BI, UNITED	
FEATURES			KINGDOM Location/Qualifiers	
source			1..1665	


```
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
1. .1665
/gene="nAcRalpha-30D"
44. .1528
/codon_start=1
/product="nicotinic acetylcholine receptor subunit
Daiphae"
/protein_id="CAD86935.1"
/db_xref="GI:29466435"
/translation="MDSPLPASLSFLVLLIFLAIKESCOGPKHEKRLNHLISLYNTL
ERPANSEPLEVKEFGLTLOQI IDVDEKNQLLINLWLSWYNDYNLRWNETYGVK
DLRITPNKLPVLMYNSADEGPDGTHSVVVGSGSCLYVPPGFKPSTCKMDITW
FPDDQHEKMFQSWTDYGNQDLVLSSEDDGSDSDFITNGEWYLLAMPKQNTIYA
CCPEYVDITFTIQRIRTYLFFNLIVPCVLISMMALLGFTLPDPSGEKLITGVTL
LSUTFLNLVAESPPTSDAVPLIGTVFNCIMFVASSVLTVVLNHYHRTADIDEM
PWTIKSVFQWLPLWLRMRGRKIRTKTILLNRMKELEKERSKSLLANVLIDID
DFRHTISGQTAIGSSASRGRPTVEEHHTAIGCNHKLHLHLILKELOFITARMKADD
EALIGDKWFAAMVDRFLIVFTLFTIIIAITVIVLLSAPHLIIVQ"
44. .121
/sig_peptide
122. .1525
/gene="nAcRalpha-30D"
/product="nicotinic acetylcholine receptor subunit
Daiphae"
ORIGIN
Query Match 49.8%; Score 747.8; DB 3; Length 1665;
Best Local Similarity 69.9%; Pred. No. 1e-187;
Matches 1044; Conservative 0; Mismatches 407; Indels 42; Gaps 1;
QY 11 TGTGGCGGCTTGGCGCTGCTGCTTTGCTGCGCCGATCGAGCAAGGTCCTCACGAGA 70
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
75 TGTGTGCTGCTGATCTTTCTGGCGATAATTAAGAAAGAGCTGTCAAGGACCTCATGAAA 134
QY 71 AGAGACTCCTGAACGCGTCTGTCGGAACACTACACACCCCTGGAGCGCCGTCGCAAG 130
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
135 AGCGCTGCTGAACCACTGCTGTCTCCACCTACAAATACGCTGGAGCGACCGCTGGCCAA 194
QY 131 AGAGCGAACCGCTAGAGCTCAGGTTTCGGCTTGACCTTGACGCAAAATCAITGACGTGACG 190
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
195 AATCGGAGCCCTGGAGGTTAAGTTTCGACCTGACGCTGCGACGAGATCATCGAGTGGACG 254
QY 191 AGAAGATCAACTTATTAACCAATATATAGCTGTCTGTTGGAGTGAATGACTTACAACC 250
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
255 AGAAGATCAGCTTCTCATACGAATCTTTGGCTTTCTGTTGGAGTGAACGACTTACAATC 314
QY 251 TGAGGTGAACGACGACGAGTATGGCGGGTCAAGGACCTCAGGATCACGCCCAACAGT 310
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
315 TGGCTGGAATGAACCGGAATACGCGGGGTCAAGGATCTACGAATCACGCCCAACAGC 374
QY 311 TGTGGAACCGGAGCTCCTTATGTATATAGTGTGACGAGGGTTTTGACGGGACCTAAC 370
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
375 TGTGGAACCGGAGCTCTCATGTACACAGCGGGATGAGGATTCGATGGCACGCTATC 434
QY 371 AGACCAACGCTGCTGTCAGAACGGCGGAGTTCCTGTACGTGCCACCTGGCATATTC 430
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
435 ACACGAGCTGTGGTCAACATGGCGGCGAGTTGCTGTACGTGCTGCCCTGGTATCTTCA 494
QY 431 AGAGCACATCAAGATGGACATCGGTGTTTCCCTTCGACGACCAACACTGTGATATGA 490
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
495 AGAGCACATCAAGATGGACATCACGTGTTTCCCTTCGATGACCAACATTTGGGAATGA 554
QY 491 AGTTGGTAGCTGGACATATGACGGCAATCAGTTGGATCTGGTGTCTAAAGATGAGGACG 550
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
555 AATTGGTAGTTGACATTCAGATGGAATCAGTTGGATTTGGTTTGGTTCCGAGATG 614
QY 551 GCGCGAGTCTATCGGACTTCATACCAATGGGGAGTGGTATCTAATAGGAATGCCAGCA 610
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
615 GAGGGGATCTTCCGATTTTCATAACAAATGGGAGTGGTACTTGTCTGGCCATGCCGGAA 674
```

REFERENCE

1 (bases 1 to 1699)
 AUTHORS
 Stapleton, M., Brokstein, P., Hong, L., Achavani, A., Carlson, J.,
 Champe, M., Chavez, C., Dorsett, V., Dresnek, D., Farfan, D., Frise, E.,
 George, R., Gonzalez, M., Guarini, H., Krommiller, B., Li, P., Liao, G.,
 Miranda, A., Mungall, C. J., Nunoo, J., Pacleb, J., Paragas, V., Park, S.,
 Patel, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Rubin, G. M.
 and Celnikier, S.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (19-DEC-2003) Berkeley Drosophila Genome Project,
 Lawrence Berkeley National Laboratory, One Cyclotron Road,
 Berkeley, CA 94720, USA
 COMMENT
 Sequence submitted by:
 Berkeley Drosophila Genome Project
 Lawrence Berkeley National Laboratory
 Berkeley, CA 94720

This clone was sequenced as part of a high-throughput process to
 sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
 Science 2000). The sequence has been subjected to integrity checks
 for sequence accuracy, presence of a polyA tail and contiguity
 within 100 kb in the genome. Thus we believe the sequence to
 reflect accurately this particular cDNA clone. However, there are
 artifacts associated with the generation of cDNA clones that may
 have not been detected in our initial analyses such as internal
 priming, priming from contaminating genomic DNA, retained introns
 due to reverse transcription of unspliced precursor RNAs, and
 reverse transcriptase errors that result in single base changes.
 For further information about this sequence, including its location
 and relationship to other sequences, please visit our Web site
 (<http://fruitfly.berkeley.edu>) or send email to
cdna@fruitfly.berkeley.edu.

FEATURES

source
 Location/Qualifiers
 1..1699

/organism="Drosophila melanogaster"

/mol_type="mRNA"

/strain="Y; cn bw sp"

/db_xref="taxon:7227"

<1..1699

/gene="nAcRalpha-30D"

/db_xref="FLYBASE:FBgn0032151"

<1..1528

/gene="nAcRalpha-30D"

/note="Longest ORF"

/codon_start=2

/product="GH15518p"

/protein_id="AAR2815.1"

/db_xref="GI:40216008"

/db_xref="FLYBASE:FBgn0032151"

/translation="PFNPRATRYKQAKMDSPASLSLFLVLLIFLAIKESCGPHE
 KRLNLHLLSTYNTLREVPANSEFLVKFGLTLQIIDVDEKNQLLIITNLWLSLEND
 YNLRWNETYGVGKDLRIIPNKLWKPDMYNSADEGPDGYHTSVVVKHGGSLYVP
 PGIFKSTCKMDITWFPDQHCMEKFGSWTDGQDLVLSDEGDLSDFITNGEY
 LLAMPKNTIYACCPPEYVDITFTQIRERTLYYFENLIVPCVLISMSALLGLTLP
 PDSEKLTGLTILLSLTVFNLVAESPTTSDAVPLIGTFNCLIMENAVSVLTVV
 VLNHRTADLHEMPPIKSVFLQWPLWLRMPGRKAITKILLISNRKLELKER
 SSKLLANVLDDIDFRHTISGQTATGSSASFGRTPTVEBHHTAIGNKHDLHLIK
 ELQFITARMRXADDEALIGDKFAAMVVDRECLIVFTLFTIITATVTVLLSAPHIIV
 "

ORIGIN

Query Match 49.8%; Score 747.8; DB 3; Length 1699;

Best Local Similarity 69.9%; Pred. No. 1e-187;

Matches 1044; Conservative 0; Mismatches 407; Indels 42; Gaps 1;

QY 11 TGTTCGGCGCCTTGGCGTGTGGCTTTGCTGCCCGTATCGAGCAAGGTCTCTCAGGAGA 70
 Db 75 TGTTCCTGTTGATCTTCTGGCGATAATTAAGAAGCTGTCAAGACCTCATGAA 134
 QY 71 AGAGACTCTGAACCGTGTGGCGAACTCAACACCCCTGGAGCGACCGTGGCCAACG 130
 Db 135 AGCGCTGTGAACCATCTGCTGTCCACTCAATACGCTGGAGGACCGGTGGCCAATG 194
 QY 131 AGAGGGAACCGTAGAGGTACGTTCCGCTTGCACCTTCAGCAATCATTCAGCTGGAGC 190

Db 195 AATCGGAGCCCTGGAGGTTAAGTTCGAGCTGCAGCAGATCATCGACGTGGACG 254
 QY 191 AGAAGATCAACTACTTATACCAATATATGGCTGTCTGTTGAGTGAATCACTACAACC 250
 Db 255 AGAAGATCACTTCTCATACGAATCTTGGCTTTTGTGGAGTGGAAACGACTACAATC 314
 QY 251 TGAGGTGGAACGACGAGGATATGGCGGGGTCAAGGACCTCAGGATCACGCCCAACAAGT 310
 Db 315 TGCCTCGAATGAACCGAATACGCGGGGTCAAGGATCTACGAATCACGCCCAACAAGC 374
 QY 311 TGTGGAAGCGGACGCTCTTATATATATAGTGTGAGGAGGTTTTCACGGGACCTTACC 370
 Db 375 TGTGGAAGCGGACGCTCTCATGTACAAACAGCGGGATGAGGGATTTCGATGGCAATC 434
 QY 371 AGACCAACGTGTGTCTCAGAAAGCGGCGAGTTCCTGTGTAGTGCACCTCGCATATCA 430
 Db 435 ACACGAGTGTGTGTCAACATGCGGCGAGTGTCTGTAGTGCCTCCCTGGTATCTTCA 494
 QY 431 AGAGCATGCAAGATGACATCCGCTGGTTTCCCTTCGACGACCAACACTGTGATAGA 490
 Db 495 AGAGCATGCAAGATGACATCACGTGGTTCCTCATTTGATGACCAACATTCGCAATGA 554
 QY 491 AGTTCCGTAGTGCACATATGACGGCAATCAGTTGGATCTGCTGTCTAAAAGATGAGCAG 550
 Db 555 AATTCCGTAGTGGACTTACGATGGAATCAGTTGGATTGGTTTGGTTTCCGGAAGATG 614
 QY 551 CGCGCGATCTATCGGACTTCATAACAATGGGAGTGGTATCTAATAGGAATGCCAGCA 610
 Db 615 GAGGGATCTTCCGATTTCAACAATGCGAGTGGTACTTGTCTGCAATGCCGGAA 674
 QY 611 AAAAGAACAAATAACATACGCTGTGCTGCCCGAGCCCTCAGTGGAGCTCACTTCAACA 670
 Db 675 AGAAGATACGATAGTCTAGCCTGTGCTGCCAGAACCATATGTGGATACACCTTACTA 734
 QY 671 TCATGATAAGAGAGCAACCTTGTACTTCTTCAACCTCATGCTCCCGTGCCTGCTGA 730
 Db 735 TACAAATTCGTCGCCGTACATTATATATTTTCAATTTAATTTGTGCATGTGTGCTTAA 794
 QY 731 TCTCATCGATGGCACTCCTCGGCTTCACTGCCACACAGACTCCGGGAGAGAACTCAAC 790
 Db 795 TCTCATCGATGGCCCTACTGGGCTTCACTTGGCCCGGATTCGGCGGAGAACTGACGC 854
 QY 791 TTGAGTCACTATTCTTCTATCGCTGACGGTTCCTCAACTGGTGTAGCGAGACCTTGC 850
 Db 855 TGGCGCTAACTATACTACTATCAATTAACAGTATTTCTAAACCTTGTTCGCCGAGTCCATGC 914
 QY 851 CACAGTCTCCGACCTATCCCCTGTTAGGAGCGTACTTCAATTTGCATCATGTTCTATGG 910
 Db 915 CGAACAACGTCGGATGCTGTTCTCTTATAGGCACCTACTTCAATTTGCATCATGTTCTATGG 974
 QY 911 TAGCGTCTCTGTGTACTGACTGTGTGTGTACTCAATTAACCACTCGAACAGCTGATA 970
 Db 975 TCGCTCTGTCGGTGTGACAGTAGTGTGTCTCACTACCACTACCACTCGACAGCGACA 1034
 QY 971 TACATGAATGCCACAGTGGATAAATCAGTATTTCTTAATGTTTGCATGATGATGCTGC 1030
 Db 1035 TTCAGGATGCCACCGCTGGATCAAGTCCGTTTCTCAATGCGTGCCTCGATCTTGC 1094
 QY 1031 GAATGTCGAGGCGAGGGAAGAAGATCACCGAAGACTATATATGATGAACAGAGATGA 1090
 Db 1095 GAATGGGTGACCCGGTCGCAAGATTACCGCAAAACAAATACTATTAGCAATCGCATGA 1154
 QY 1091 GGGAGCTGGAACCTGAAGAGAGGTGTCGAAGTCTTCTGCTGGCGAATCTTCTAGATATG 1150
 Db 1155 AGGAGCTGGAGCTAAAGGAGCGCTCTCCAAATCCCTGCTGGCCAATGTCTCTCGACATCG 1214
 QY 1151 ATGATGACTTCAGACACGCGCCCTCGCCTCTCAACAGTACTGCCTCGACCGGGAATTGG 1210
 Db 1215 ACGACGACTTCGGGCACACAATATCTG----- 1241
 QY 1211 GACCTGGGTGCTCAATATTCGCGACGGAATTTCCGTCGGTGTTCGTCGCTCCGTCACGGA 1270
 Db 1242 -----GCTCCCAACGCCCATTTGGCTCGTCGGCCAGCTTCGTCGGCCCAACAACGG 1292

830 TCAACGCTAGCGGAGTGTCTGTACGTGCCCTGGTATCTTCAAGAGCATGCAAGA 889
446 TGGACATCGGTGTTCCCTTCGACGACCAACTGTGATATGAAGTTCGGTAGCTGGA 505
890 TGGACATCAGCTGTTCCCTTCGACGACCAACTGTGATATGAAGTTCGGTAGCTGGA 949
506 CATATGACGGCAATCAGTTGATCTGTGCTTAAAGATGAGCGAGGCGGCATCTATCGG 565
950 CTTAGATGGAATCAGTTGATCTGTGCTTAAAGATGAGCGAGGCGGCATCTATCGG 1009
566 ACTTCATAAATAAGGAGTGTATCTAATAGGAATGCCAGGCAAAAGAACACATAA 625
1010 ATTTATAAATAAGGAGTGTATCTAATAGGAATGCCAGGCAAAAGAACACATAA 1069
626 CATAGCGGTGCTGCCCGAGCCCTACGCGGAGCTCACCTTCACCATCATGATAGAGAC 685
1070 TCTAGCGCTGCTGCCCGAGCCCTACGCGGAGCTCACCTTCACCATCATGATAGAGAC 1129
686 GAACTCTGTACTCTTCAACCTGTGCTGCCGTGCTGCTCATCTCATGATGGCAC 745
1130 GTACATATATATTTTCAATTTAATCGTGCCATGTGCTAATCTCATGATGGCCC 1189
746 TCCTCGGCTTACACTGCGACAGACTCGGAGAGAACTCACACTTGGAGTCACTATTC 805
1190 TACTGGGCTTACACTGCGGAGTTCGGGCGAAACTGACGCTGGGAGTTACAATTC 1249
806 TTCTATCGCTGACGCTGTTCTCAACCTGTGAGCGAGACCTGCGACAGTCTCCGACG 865
1250 TTCTATCGCTGACGCTGTTCTCAACCTGTGAGCTGAGACATGCCCCAAGTACTGATG 1309
866 CTATCCCCCTGTTAGGAGCTACTTCAATTCGATCATGTTTATGCTAGCGTCTGCTGG 925
1310 CAATCCCCCTGTTAGGAGCTACTTCAATTCGATCATGTTTATGCTAGCGTCTGCTGG 1369
926 TACTGACTGTGCTGATCAATTAACCACTGAGCAAGCTGATATATGAAATGCCAC 985
1370 TGCTGACAGTATGCTGCTCAACTACCACTGAGCAAGCTGATATGAAATGCCAC 1429
986 AGTGATATAATCACTATTCTTACAAATGCTTCCCATGATATGCGAATGTCGAGCCAG 1045
1430 GGTGATCAAGTCCGTTTCTTCAATGCTGCTGCTGCTGATCTTGGATGCTGACCCG 1489
1046 GGAAGAAGATCACGAGGAGCTATATGATGAACACGAGGATGAGGAGCTGGAATGGA 1105
1490 GTGCAAGATTACACGCAAAACAATATTAAGCAATCGCATGAAGGAGCTGGAGCTAA 1549
1106 AGGAGGCTGCTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1165
1550 AGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1609
1166 ACGGCCCTCCGCTCTTCAACAGTACTGCTCGACCGGAAATTTGGGACCTGGGTGCTCAA 1225
1610 ACAATATCTG-----GCTCC 1627
1226 TATTCGCGCAGGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1285
1628 AAACCGCATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1687
1286 GCGGCTGGTAGCACCATCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1345
1688 CGGCCATCGGCTGCAATCAAAAGATCTTCAATTAATTTTAAAGAAATTTGCAATTTATTA 1747
1346 CGGCCAGGATGAAGAGCTGATGAGGAAGCGGAGCTGATCAGCGACTGGAAAGTTGCTG 1405
1748 CGGCGGAGTGCAGAGCTGACGAGCGAAGCGAATGATCGCGATTGGAAGTTGCGCG 1807
1406 CGATGTTGTTGATAGGTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1465
1808 CAATGTTGTTGATAGATTTGTTTAAATGTTTCACTGCTGCTGCTGCTGCTGCTGCTGCT 1867
1466 TAGCTGCTGCTGTTATCGGACCGCATATCATCTGTCGAA 1503
1868 TTACGTTGCTGCTCTCCGCTCCGACATATGTCGAA 1905

RESULT 10
AF321449
LOCUS
DEFINITION
2110 bp mRNA linear INV 29-APR-2002
Drosophila melanogaster nicotinic acetylcholine receptor Dalpha6
subunit variant type V (nAcRalpha-30D) mRNA, complete cds,
alternatively spliced.
ACCESSION
AF321449
VERSION
AF321449.1 GI:20152852
KEYWORDS
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 2110)
Grauso, M., Keenan, R.A., Culetto, E. and Sattelle, D.B.
Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
Dalpha5, Dalpha6 and Dalpha7, in Drosophila melanogaster Identify a
New and Highly Conserved Target of Adenosine Deaminase Acting on
RNA-Mediated A-to-I Pre-mRNA Editing
Genetics 160 (4), 1519-1533 (2002)
21969411
MEDLINE
11973307
PUBMED
REFERENCE
2 (bases 1 to 2110)
Grauso, M. and Sattelle, D.B.
Direct Submission
Submitted (15-NOV-2000) Human Anatomy and Genetics-Functional
Genetics Unit, University of Oxford, South Park Road, Oxford OX1
3QX, UK
FEATURES
source
1. .2110
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/chromosome="2"
/map="30D1"
/dev_stage="embryo"
1. .2110
/gene="nAcRalpha-30D"
379..1950
/gene="nAcRalpha-30D"
/note="ion channel; neurotransmitter transmembrane
receptor; alternatively spliced; contains exons 3a, 8a and
8b; contains two repeats of the transmembrane TM2 region"
/codon_start=1
/product="nicotinic acetylcholine receptor Dalpha6 subunit
variant type V"
/protein_id="AA03396.1"
/db_xref="GI:20152853"
/translation="MDSPLASLFLVLLIFLAIKESCGQPHKRLNHLSTYNTL
BRPVANESEPLEVFGFLTQOI IDVEKNOILTNAWLNLEWNDYLNWNETEGGVK
DLRTPNKLKPDVLMYNSADGEGDVTHTNVKNGSCLYVPPGIFKSTCKMDITW
FPDQHCMEKFGWYDCNQIDLVNSEDGDLSDFTNGEWILLAMPKNTIYA
CCPEYVDITTIQIRRTLYFFNLIVPCVLISSMALLGFTLPDPSGEKLTGLVTIL
LSLTVNLVAESMPTTSDAVPLIGTITLLSLTVNLVAETLPQVSDAIPULGYFN
CTMFVASSVLTIVVNLNHRHTADHEMPPIKISVFLQWLPWILMRMGPKRTKRT
ILLSNRKELBELKSSKLLANLVIDDDFRHTISGQTGSAIGSASFGRPTTVEHH
TAJGNHKDLHLIKELQFITARMKADDEALIGDWKFAAMVVDVDFCLIVTFLTII
ATVVLSSAPHLIVQ"
846
/gene="nAcRalpha-30D"
/note="results in methionine to isoleucine substitution;
compared to the sequence deposited in GenBank Accession
Number AE003626"
/replace="a"
1401
/gene="nAcRalpha-30D"
/note="compared to the sequence deposited in GenBank
Accession Number AE003626"
/replace="c"
1647
variation
variation
variation

TITLE Direct Submission
JOURNAL Submitted (19-APR-1999) ZF-BTB, Bayer AG, Bldg. Q 18,, Leverkusen
51368, Germany

FEATURES
source Location/Qualifiers
1..3629
/organism="Heliothis virescens"
/mol_type="mRNA"
/db_xref="taxon:7102"
335..1825
/note="hvnachra7-1"
/codon_start=1
/product="putative nicotinic acetylcholine receptor alpha
7-1 subunit"
/protein_id="AAD32697.1"
/db_xref="GI:4895005"
translators="MGGARRSHLAAPAGLLLLCLLWPRGARQGYHEKRLHLLLDH
YNYLERPVNESDPLQSLFGLTLMQIIDVDEKQOLLITNWLKLEWDMNLRWNTSDF
GGVKDLRVPRLHRLMDVLTNSADFGSDTPTNVVVRNNGSLYYPGPFKSTCKI
DIYPPDDORCEKMFQSWYDGYLDLODEGGDISFVTVNGEWELI.GVPGKRNE
IYVNCPEYDITFAVIRKTLTYFNLIVPCVLIASMALLGFTLPDPSGEKLSLG
VTILSLTVELNMVAETMPGSDAVPLGLTYFNCIMFWVASVSVSTILLNLYHHRHAD
THEMSDWIRCVELVWLPWLRMSRPSGATPPPARVPPDLELRERSSKILLANVLD
IDDFRHPQAOQPCRCYRYRGGENGAGLAHSCFQVDYELSLILKEIRVITDQMRKD
DEDADISRDKWPFAMVDRLLIIFTLTIATLAVLLSAPHINVS"

ORIGIN

Query Match 38.3%; Score 575; DB 3; Length 3629;
Best Local Similarity 64.6%; Pred. No. 1.1e-141;
Matches 929; Conservative 0; Mismatches 460; Indels 48; Gaps 3;

Qy 64 CACGAGAGAGACTCTGAAACCGCTGCTGGCGAACTACAAACCCCTGGAGCGACCGGTG 123
Db 431 CACGAGAGCGCTACTGCACCACTATTGGACCACTACAAACGCTACTGGAGAGCGCGTC 490

Qy 124 GCCAACGAGACGAAACCGCTAGAGTCAGGTTCCGCTTGACCTTGCAGCAATCATTTGAC 183
Db 491 GTCAACGAGACGACCGCTGCAGCTCTCTTCGGCCTCAGCTCATGTCAGATCATCGAC 550

Qy 184 GTGACGAGAGAAATCAACTACTTATAACCAATATATGGCTGTCTGTTGGAGTGAATGAC 243
Db 551 GTGACGAGAGAACCGACTTTTATAACCAACATCTGGCTAAACCTAGAGTGAATGAT 610

Qy 244 TACAACTGAGGTGGAACGACGAGGATATGGCGGGGTCAAGGACCTCAGGATCAAGCCC 303
Db 611 ATGAACCTGAGGTGGAACACTTTCAGATTTCGCGCGGGGTCAAGATTTAAGAGTGCACCC 670

Qy 304 AACAGTGTGGAACCGGAGCTCTTATGATATAGTCTGACGAGGGTTTTCAGCGG 363
Db 671 CACAGACTATGGAACACGAGCTCTTATGATACAAACGCGGAGGAGGGTTTCGACAGC 730

Qy 364 ACCTACCAGACCAACGTTGGTTCAGAGCGCGGCGAGTTGCCCTGTAGTGCACCTGGC 423
Db 731 ACGTATCCAAACGAGTGTGTGCGGAACAAACGGCTCGTGTCTGTAGTGCACCGCGC 790

Qy 424 ATATTCAAGACCATGCAAGATGGAACATCGCGGTGGTTTCCCTTCGAGACCAACACTGT 483
Db 791 ATCTTCAAGACCATGCAAGATGGAACATCACTCGTTTCCCTTCGAGACCAACAGTGC 850

Qy 484 GATATGAATTCGCTAGCTGACATATGACCGCAATCAGTTGGATCTGCTTAAAGAT 543
Db 851 GAGATGAATTTGGCAGCTGGACTTATGATGTTTATCAGTTGGATCTCAAACTACAGAT 910

Qy 544 GAGCAGCGCGGATCTATCGACTTTCATAACAAATGGGAGTGGTATCTTAATAGGAATG 603
Db 911 GAACGGGCGGAGATATAAGCAGTTTGTACGAATGCGAATGGAGTTAATAGGATC 970

Qy 604 CCAGGCAAAAGAACACAAATACATACCGCTGCTGCCCGAGCCCTACGTGGAGTCAC 663
Db 971 CCCGGCAAGCGCAACGAGATCTACTACAACTGTTGTCCGAGGCCATACATCGATCACG 1030

Qy 664 TTCACCATCATGATTAAGAGACGACCTTGTACTTCTTCACTTCACTGATCGTCCGTCG 723
Db 1031 TTTGCGGTGTTGATCCGAGGAAACGCTCTACTTCTTCACTGATCGTCCGTCG 1090

Qy 724 GTGCTGATCTCATCGATGGCACTCCTCGGCTTTCACACTGCCACGAGACTCCGGAGAGAA 783
Db 1091 GTGCTCATCGCTCCATGGCTCTATTGGGGTTTACCTTGGCTCCAGACTCCGGAGAAAG 1150

Qy 784 CTCACACTTGGAGTCACTATTCTTCTATCGCTGACGGTGTTCCTCAACCTGGTACCGAG 843
Db 1151 TTGCTTTAGTGTGACGATATTACTGTCGTGACGGTGTTCCTCAACATGGTGGCGGAG 1210

Qy 844 ACCCTGCCACAGGTCTCCGACGCTATCCCTGTTAGGGAGTACTTCAATTTGATCATG 903
Db 1211 ACGATGCCAGGACGTCGGACCGCTGCTCCCTTGTCTCGGCACTTCAACTGATCATG 1270

Qy 904 TTCATGTTAGTCTGCTGTGGTACTGACTGTGTGGTGTACTCAATTAACCACTCGAACA 963
Db 1271 TTCATGTTGGCTTCTCCGTCGTCTCCACCATCTGATCTCTCAACTACCACCGGCAC 1330

Qy 964 GCTGATATCATGAAATGCCACAGTGATGATAAATCAGTATTTCTTACAAATGGTTGCATGG 1023
Db 1331 GCAGACACTCACGAAATGAGTGTGTTGCGTGTTCCTTTATTTGGCTGCCGTGG 1390

Qy 1024 ATACTGCGAATGTCGAGCGCCAGGAGAAAGATCACACGAGAACTATAATGATGAACACG 1083
Db 1391 GTGCTGCGATGTACGCGCCGCTCGCG---ACGACGCGCGCGCGCGGCTACCT 1447

Qy 1084 AGGATGAGGAGCTGGAACCTGAAGGAGAGGTCTGTCGAAGTCTTGTGCGGAATGTTCTA 1143
Db 1448 CCGCGCGGACCTGGAGCTGCGGAGCGCTCTCTCCAAGTCGCTCTTAGCGAACGTGCTC 1507

Qy 1144 GATATGATGATGATCTTACAGACGCGCTTCGCGCTCTTACACATCTGCTCGACCGGG 1203
Db 1508 GACATGATGACGACTTTCGCGCACCCGCAAGCGAGCAG-----CCGCAATGCTGCCG 1560

Qy 1204 AATTGGGACCTGGGTCTCAATATCCGACGAGATTCGTCGTCGTTGTCGTCGTCG 1263
Db 1561 ATACTACAGGGGGGTGAGGAGAAATGCGCGGGTTGGCGCGCACAGTTGCTTCG---- 1616

Qy 1264 TCCACGATGAAGACGTGGCGGGCTGGGTAGCCACCACCATCCGAGCTGCACTCAT 1323
Db 1617 -----GTGTCGACTACGAGCTCTCCCTCAT 1642

Qy 1324 CTGAGAGCTGCGATTTCATCAGCGCCAGGATGAAGAGGCTGATGAGGAGCGGAGTGG 1383
Db 1643 CTGAAGAGATATAGAGTCATCAGATCAGATGCGCAAGGACGACGAGATGCGGACAT 1702

Qy 1384 ATCAGCACTGGAAGTTTGTGCGATGCTGTTGATAGTGTTCCTGTTGCTGTTTCA 1443
Db 1703 TCGCGCACTGGAAGTTTGGCGCCATGCTGCTGAGACAGACTGTGCTTATTTTACC 1762

Qy 1444 CTTTTCACATCATCGGACAGTAGTGTCTCTGTTATCGGACCGCATATCATCGT 1500
Db 1763 CTGTTCAATCATCGCACGCTAGCCGTGCTGCTGTCGCGGCGCACACATCATGGT 1819

RESULT 12

AX009612
LOCUS 3700 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 3 from Patent EP0962528.
ACCESSION AX009612
VERSION AX009612.1 GI:9996844
KEYWORDS
SOURCE Heliothis virescens (tobacco budworm)
ORGANISM Heliothis virescens
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuoidea; Noctuidae; Heliothinae; Heliothis.
1
REFERENCE
AUTHORS Adamczewski M.D., Schulte, T.D. and Oellers, N.D.
TITLE Nucleic acids encoding acetylcholin-receptor subunits from insects
JOURNAL Patent: EP 0962528-A 3 08-DEC-1999;
BAYER AG (DE)
FEATURES
source Location/Qualifiers
1..3700

```

/db xref="GI:29466437"

```


Genetics 160 (4), 1519-1533 (2002)	JOURNAL
21969411	MEDLINE
11973307	PUBMED
2 (bases 1 to 2907)	REFERENCE
Grauso.M. and Sattelle,D.B.	AUTHORS
Direct Submission	TITLE
Submitted (27-MAY-2000) Human Anatomy and Genetics, MRC-FGU, University of Oxford, South Park Road, Oxford OX1 3QX, UK	JOURNAL
Location/Qualifiers	
1. .2907	FEATURES
source	
/organism="Drosophila melanogaster"	
/mol_type="mRNA"	
/db_xref="taxon:7227"	
/chromosome="2"	
/map="2L; 34E4-5"	
/dev_stage="embryo"	
1. .2907	gene
/gene="nAcRalpha-34E"	
/allele="A"	
289. .2712	CDS
/gene="nAcRalpha-34E"	
/note="ion channel; neurotransmitter transmembrane receptor"	
/codon_start=1	
/product="nicotinic acetylcholine receptor Dalpha5 subunit"	
/protein_id="AA013390.1"	
/db_xref="GI:20152840"	
/translation="MKNAQLKTEVDDDELAVLRHLCNSNFSSSSSTRTSSNQRH NQLSTLPLRSLSLTKHNSIASEQHNSQQQEPASKEDEVANHGSRNDQPHLQQLDSS NMLSPKTAATAAGDEATTOQPTNIRLCARQRRLRRRKRKPATPNEDTIKKQQL SMPPTKSTDTYTPAATSCPTATQCRASDNFSPISRHORVSTATFAWVLH VLOVLVLSQWLHVQORSVLLPFRRAATAFISVYLSGFAAQLKNSSSSSSSNS NNSSTQILNGKHSWIPLLIVNLKAVCLAGVHKRLHLLDLPVNTLERVLNES DPLQSPGLTLMQIDVDENQLVLTVNLKLEWMDNLRWNTSDYGVKDLRIPHR DPLQVLMINGADEGFYQTNVVRNNGSLVPPGIFKSTCKIDITWFPDDQRC EMKFSWYDFQDLQDQSTGDISVYNGEWELLGPGKRNEIYNCCPEPID ITFAIIIRRLTYFFNNLIIPCVLIASMALIGFTLPDPSGEKLSLIGTILLSLTFIN LVAETMPATSDAVPLGTYFCNFMVASSVSTILINVHNRNADTHESEWIRIVF MCWLPWILMRSPGRPLILEPTTPCSDTSSRKHILSDVELKERSKSLANVLDI DDDPHNCRPMTPGTLPHNAPFRTYVYGGDDSGIGPIGSTPMFPAVHTHCIKST EYELGLILKEIRFTDQKRDCEANDMKFAMVVDRLCLLIIFTMFAILLATIAVL LSAPHILIVS"	
373	variation
/gene="nAcRalpha-34E"	
/note="compared to B allele"	
/replace="a"	
375	variation
/gene="nAcRalpha-34E"	
/note="results in phenylalanine to isoleucine substitution; compared to B allele"	
/replace="c"	
827	variation
/gene="nAcRalpha-34E"	
/note="results in threonine to isoleucine substitution; compared to B allele"	
/replace="t"	
1079. .1080	variation
/gene="nAcRalpha-34E"	
/note="results in lysine to arginine substitution; compared to B allele"	
/replace="gg"	
1081. .1086	variation
/gene="nAcRalpha-34E"	
/note="results in deletion of asparagine and serine; compared to B allele"	
/replace=""	
1119*1120	variation
/gene="nAcRalpha-34E"	
/note="results in serine, glycine, and asparagine insertion; compared to B allele"	
/replace="agcggcaac"	
1121	variation
/gene="nAcRalpha-34E"	
/note="results in asparagine to serine substitution; compared to B allele"	
/replace="g"	
ORIGIN	
Query Match 34.7%; Score 521.6; DB 3; Length 2907; Best Local Similarity 61.8%; Pred. No. 1.9e-127; Matches 924; Conservative 0; Mismatches 519; Indels 51; Gaps 4;	
QY 57 AGGTCTCTCAGAGAAGAGACTCTTGAACGGTCTGTCGCGAACTTACAAACCCCTGAGCG 116	
DB 1212 AGGATATCATGAAGAAGACTGTTCACGATCTTTTGGATCCTTATTAATACTAGAACG 1271	
QY 117 ACCGTTGCCAAACGAGACGCAACCGCTAGAGGTCAAGTTCAGGTTCGGCTTGACCTTGACGCAAAAT 176	
DB 1272 TCCCGTTCTCAATGAATCGGACCCGTTACAATTAAGCTTTTGGTTAACTTTAATTCGAAAT 1331	
QY 177 CATTGACCTGGACGAGAAGAAATCAACTACTTATTAACCAATATATGGCTGTCGTTGGAGTG 236	
DB 1332 TATCGATGTGGACGAGAAAAATCAATTGCTAGTCATTAATGTGTGGTTAAAACTGGAGTG 1391	
QY 237 GAATCACTACAACCTGAGGTGGAAACGACGACGAGTATGGCGGGGTCAAGGACCTCAGGAT 296	
DB 1392 GAAACGACATGAATCTCCGTCGGAACACTCCGACTATGGCGGAGTTAAGGATCTCGGAAT 1451	
QY 297 CACGCCCAACAAGTTGTGGAAAGCGGACGCTCTTATGTATATAATAGTGCTACAGAGGTTT 356	
DB 1452 ACCGCCGATCGCATCTGTGAAGCGGACGCTGATGTACAACAGTGGCATGAGGAT 1511	
QY 357 TGACGGGACCTTACCAGACCAACGCTGTGTGTGTCAGAAAGCGGGCGGAGTTGCTGTACGTGCC 416	
DB 1512 TGACGGCACCTTACCAGACGACGTGTGTGTGTCGGGAAACAACGGCTCGTGTCTATACGTTCC 1571	
QY 417 ACCTGGCATATTCAAGAGCATCATGCAAGATGAGACATCGCTGTGTTTCCCTTTCAGACCA 476	
DB 1572 GCGCGGGATCTTCAAGTCGACGCTGCAAGATCGACATCACTGTGTTCCTTTCGATGACCA 1631	
QY 477 ACACCTGTGATATGAAGTTTCGTGTAGCTGGACATATGACGGCAATCAGTTGGATCTGTGCT 536	
DB 1632 CGGTGCGAGATGAAGTTTCGGCAGTTGGACCTACGAGGATTCAGCTGGAATTTACAAT 1691	
QY 537 AAAAGATGAGGCGGCGGATCTTATCGACTTTCATAACAAATGGGAGTGTGTATCTAAT 596	
DB 1692 ACAAGATGAACTGGCGTGATATCAGCAGTTACGTCTCAACGGGAGTGGGAATCTACT 1751	
QY 597 AGGAATGCCAGGCAAAAAGAACAAATAACATACCGGTGCTGCCCCGAGCCCTACGTGGA 656	
DB 1752 GGGTGTGCCCGGCAACGTAACGAGATCTATTACAACCTGTCGCCGGAACCCCTATATAGA 1811	
QY 657 CGTCACTTTCACCATCATGATGAAGAAGACGAAACCTTGTACTACTTCTTCAACCTGATCGT 716	
DB 1812 CATCACCTTCGCCATCATATCCGCCGACGAAACACTGTACTATTCTTCAACCTGATCAT 1871	
QY 717 CCGTGCCTGTGATCTCATGATGGCACTTCCTGGGTTTCAACATGCAACGACACGACCTCCGG 776	
DB 1872 ACCTTGTACTGATTGCCTCCATGGCTTGTCTCGGATTCACCTGTCGCCGACAGATTCGGG 1931	
QY 777 AGAGAACTCACCTTGAGTCACTATTCTTCTATCGTCGACGGTGTCTCAACTGCT 836	
DB 1932 TGAATAATATCGTGGGTGTACCATCTTGTCTCTGCTACCGCTGTTTCTGAAATATGGT 1991	
QY 837 AGCCGAGACCTGCCACGAGTCTCCGACGCTATCCCCCTGTTTAGGACCGTACTTCAATTTG 896	
DB 1992 TGCCGAGACAAATGCCGGCTACTTCCGATGGGTGCCATGCTGGGTACATATTTCATTTG 2051	
QY 897 CATCATGTTCAATGTAGCTGCTGTGTGTACTGACTGTGTGTGTGTCTCAATTTACACCA 956	
DB 2052 CATTAATGTTATGTAGCTTCAATCCGTTGTGTCAACGATTTTAATTAATTAATATATCA 2111	
QY 957 TCGAACAGCTGATATACATGAAATGCCACAGTGGATAAATCATAGTATTCCTACATGGTT 1016	
DB 2112 TCGAAATGCTGATACGACGCAAAATGTCCGAATGGATACGATCGTGTGTTTGTGCTGGCT 2171	

QY	1017	GCATGGATACTCGGAATGTCGAGGCCAGGAAGAGATCACCAGGNAG-----	1065
Db	2172	GCATGGATATTGCGAATGAGTCGCCACGAGACCGCTGATCCTAGAGTTCCCGACCCAC	2231
QY	1066	-----ACTATAATGATGAACACGAGGATGAGGGAGCTGGAACTGAA	1106
Db	2232	GCCTGTTCCGACACATCCTCCGAGCGGAAGCACCGATACTCTCCGACGTTGAGCTGAA	2291
QY	1107	GGAGAGTCTGTCGAAGTCCCTTGTCGGCAATGTTCTAGATATTGATGATGACTTCAGACA	1166
Db	2292	AGAGCGCTCTCGAAATCGCTGTGGCCAACTGACTAGACATCGATGACTTCCGGCA	2351
QY	1167	CGGCCCTC-----CGCCTCCTAACAGTACTGCCTCGACCCGGAAATTGGGACCT	1215
Db	2352	CAATGTGCGCCCATGACGCCCGCGGAACACTGCCACAACCCCGCTTTCTATCGCAC	2411
QY	1216	GGGTGCTCAATATTCCG-CACGGATTTCCTGTCGTCGTCCTCCGTCCACGATGGA	1274
Db	2412	GGTTTATGGACAAGGCGACGATGGCAGCATTTGGCCCAATTGGCAGCACCCGAAATGCCGA	2471
QY	1275	AGACGTGGGGCGGGCTGGGTAGCCA-----CCATCGCGAGCTGCACCTCATCT	1325
Db	2472	TGCGGTACCCCATCATACGTGCATCAAAATCATCAACTGAATATAGGTTAATCTT	2531
QY	1326	GACAGAGCTGCAGTTTCATCAGCGCCAGGATGAAGAGGCTGATGAGGAAGCCGAGCTGAT	1385
Db	2532	AAAGGAAATTCGCTTTATAAATGATCAGCTACGTAAGATGACGAGTGCAATGACATTCG	2591
QY	1386	CAGCGACTGGAAGTTTGCTCGGATGGTTGTTGATAGGTTTTCGCTGTTTCGTTTCACACT	1445
Db	2592	CAATGATTGGAATTTGCACCTATGGTCGTTGACAGACTGTGCTTATCATATTCACAAT	2651
QY	1446	TTTCACAAATCATCGCGACAGTAGCTGTCCTGTTATCGGCACCGCATATCATCGT	1499
Db	2652	GTTCGCAATATTAGCCACAATAGTGTACTACTATCATCAGCACCAATATTATTGT	2705

Search completed: May 8, 2004, 02:41:32
 Job time : 5963.59 secs

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 10:08:00 ; Search time 604.489 Seconds

(without alignments)
10562.710 Million cell updates/sec

Title: US-09-303-232-5_COPY_95_1597

Perfect score: 1503

Sequence: 1 atggccctcatgtgccc.....caccgcatacatcgtgcaa 1503

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 212409041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_29Jan04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001s:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002s:*
- 7: Geneseqn2003as:*
- 8: Geneseqn2003bs:*
- 9: Geneseqn2003cs:*
- 10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1503	100.0	3109	3	Aaz24477 H. viresc
2	575.6	38.3	1540	4	Abi07231 Drosophil
3	575	38.3	3700	3	Aaz24476 H. viresc
4	430.6	28.6	936	4	Abli13733 Drosophil
5	411.4	27.4	2886	3	Aaz24475 D. melano
6	377.2	25.1	1509	3	Aac58395 Human PRO
7	374	24.9	1509	4	Aac90380 Wild-type
8	374	24.9	1876	2	Aat48239 Neuronal
9	374	24.9	1876	8	Adai10864 Human neu
10	373.6	24.9	1530	2	Aav44687 V274T var
11	373.6	24.9	1876	2	Aav12197 Human neu
12	373.6	24.9	1876	6	Abss4875 Human neu
13	373.6	24.9	1876	6	Abv73248 Human neu
14	372.4	24.8	1509	4	Aac90385
15	372.4	24.8	1509	4	Aac90386
16	370.8	24.7	1509	4	Aac90387
17	363.2	24.2	1964	6	Abz11298 Human pol
18	342.4	22.8	2769	2	Aat59196
19	326.8	21.7	2101	2	Aat59197
20	322.8	21.5	1869	6	Aal45873
21	322.8	21.5	1869	6	Abli54794 Modified
22	304.4	20.3	2082	2	Aat59527
23	304.4	20.3	2752	9	Adb78668 Human nic

24	304.4	20.3	2752	9	Adb78671 Human nic
25	304.4	20.3	2752	9	Adb78670 Human nic
26	304.4	20.3	2752	9	Adb78661 Human nic
27	304.4	20.3	2752	9	Adb78672 Human nic
28	304.4	20.3	2752	9	Adb78663 Human nic
29	304.4	20.3	2752	9	Adb78669 Human nic
30	304.4	20.3	2752	9	Adb78662 Human nic
31	304.4	20.3	2752	9	Adb78673 Human nic
32	304.4	20.3	3343	9	Ade85024 Farnesyl
33	304.4	20.3	3496	2	Aat48237 Neuronal
34	304.4	20.3	3496	8	Adai10858 Human neu
35	303.8	20.2	2666	9	ADC71168 Human neu
36	302.8	20.1	1809	5	AAS91552
37	302.8	20.1	2082	2	AAT59528 DNA encod
38	302.8	20.1	2752	9	Aas9528 Alpha4 su
39	302.2	20.1	1908	2	Adb78660 Human nic
40	302.2	20.1	1908	8	Aat48236 Neuronal
41	302.2	20.1	2277	2	Adai10856 Human neu
42	302.2	20.1	2277	2	Aav12199 Human neu
43	302.2	20.1	2277	2	Aaq90387 Alpha 2 s
44	302.2	20.1	2277	6	Abss4870 Human neu
45	302.2	20.1	2664	2	Abv73243 Human neu
					Aat48235 Neuronal

ALIGNMENTS

RESULT 1

Aaz24477

ID AAZ24477 standard; cdna to mRNA; 3109 BP.

XX

AC AAZ24477;

XX

DT 17-FEB-2000 (first entry)

XX

DE H. virescens acetyl-choline receptor DNA from clone Hva7-2.

XX

DE Acetyl-choline receptor; nicotinic; insect; insecticide; screening;
KW neurotransmission; plant protection agent; conductance; AChR; ds.

XX

OS Heliothis virescens.

XX

PN DB19819829-Al.

XX

PD 11-NOV-1999.

XX

PF 04-MAY-1998; 98DE-01019829.

XX

PR 04-MAY-1998; 98DE-01019829.

XX

PA (FARB) BAYER AG.

XX

PI Adamczewski M, Oellers N, Schulte T;

XX

DR WPI; 2000-014207/02.

XX

DR P-PSDB; AAY50816.

XX

PT New nucleic acid encoding a nicotinic acetylcholine receptor from

insects, used to identify potential insecticides.

XX

PS Claim 1a; Page 19-22; 26pp; German.

XX

CC This invention describes a novel nucleic acid (NA) encoding a nicotinic
acetyl-choline receptor (I) from insects which can be used as an
insecticide. Inhibitors of (I) interfere with neurotransmission. (I)
CC (also vectors containing it, its regulatory regions, and antibodies
directed against (I)-encoded proteins) are used to screen for: (a) plant
protection agents that alter conductance of AChR, potentially useful as
insecticides, or (b) genes which encode polypeptides that are involved in
formation of functionally related AChR in insects. (I) are also used to
isolate and characterize the specified regulatory regions and for
recombinant production of (II). This sequence encodes an acetyl-choline
receptor isolated from *Heliothis virescens*

XX Drosophila melanogaster.
OS WO200171042-A2.
PN
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.
XX
XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
PR
XX (PEKE) PE CORP NY.
PA
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
PI
XX WPI; 2001-656860/75.
DR P-PSDB; ABB69630.
DR
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
XX Claim 1; SEQ ID NO 35681; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from wipo at ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 936 BP; 238 A; 213 C; 239 G; 246 T; 0 U; 0 Other;
XX
XX Query Match 28.6%; Score 430.6; DB 4; Length 936;
XX Best Local Similarity 68.5%; Pred. No. 2.4e-109;
XX Matches 595; Conservative 0; Mismatches 274; Indels 0; Gaps 0;
XX
XX 19 GCCTTGGCGCTGTGGCTTGTCTGCGCGTATCGAGCAAGTCTCCACGAGAGAGACTC 78
XX
XX 67 GGCCTGGACTTTTAATTATGATACCGCTTGTGCGGTGGACCCCATGAGAACGGCTA 126
XX
XX 79 CTGAACGCGTGTGTGCGGAACCTACACACCTCGAGCGACCGGTGGCCCAACGAGCGAA 138
XX
XX 127 CTCACGCGCTTCTGGACAACCTACACAGCCTGGAGCGTCCGGTGCATGAATCCGAT 186
XX
XX 139 CGCTAGAGGTGAGTTCGGCTTGACCTTCAGCAAAATATTGACGTGGACGAGAGAAAT 198
XX
XX 187 CCATTGCACTGAGCTTCGGACTTAACACTCATGCAATTCAGATTATCGATGGACGAAAGAAT 246
XX
XX 199 CAACTACTTATTAACCAATATATGGCTGTCTGGAGTGAATGACTACAACTGAGGTGG 258
XX
XX 247 CAACTGCTTATACGATATTTGGCTCAATGGAATGGAACGATGATGATCTTCGATGG 306
XX
XX 259 AACGACGCGAGTATGCGGGGTCAGGACCTCAGGATCAGCCCAACAAAGTTGTGGAAG 318
XX
XX 307 AATTCGAGTGAGTTCGGTGTGCGGGATCTCGGAATTCGCGCATCTCGCCTATGAAA 366
XX
XX 319 CCGGAGCTCTTATGATATGATGCTGAGGAGGTTTTCGCGGACCTTACGAGCAAC 378
XX
XX 367 CCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 426
XX
XX 379 GTGGTGGTCAGAACGCGCGGCTGCTGCTGATGCTGCCACTTGGCATATTCAGAGCA 438
XX
XX 427 GTGGTGGTTCGCAATTAATGGAGCTGTCTGTACCTACCGCGAGTATATTTAGTCAACG 486
XX
XX 439 TGAAGATGAGATCCCGTGGTTCCTTCGACGACCAACACTGTGATGATGAAGTTTCGTT 498
XX
XX 487 TGTAAGATCGACATTACGTGGTTCCTTCGATTCGATTCGATTCGATTCGATTCGATTCG 546

QY 499 AGCTGACATATGACGGCAATCAGTTGGATCTGGTCTAAAAGATGAGGAGCGCGCAT 558
Db
QY 547 TCGTGGACCTTACGATGGGTTTCAGTTGGACCTTCAGTTGGACGAAAGCTGGTGGCGAC 606
Db
QY 559 CTATCGGACTTCATAACAATGGGAGTGGTATCTAATAGGAATGCCAGCAAAAGAAC 618
Db
QY 607 ATTCTAGCTTTATAACCAATGGGAATGGACTTGTAGGTGTGCCCGGTAAACGAAAT 666
QY 619 ACAATAACATACGCTGTCTGCCCGGAGCCCTACGTGGACGTCACCTTCCACATCATGATA 678
Db
QY 667 GAAATCTACTATAATTTGCTGCCAGAACCTTATATTGACATACATTCGCCATTTTGATA 726
QY 679 AGAAGACGAACCTTGTAATCTTCTTCAACCTGATCGTCCCGTGGTGTGATCTCATCG 738
Db
QY 727 AGCGCAAAACGTTGTACTATTTTCAATCTGATTTGCGGTGCGTACTGATCGCGCTCC 786
QY 739 ATGGCACTCTCGGCTTCACTGCGGAGCTCCGAGAGAACTCACACTTGGAGTC 798
Db
QY 787 ATGGCACTCTGAGGTTTACCTGCCACGAGATCTGGTGAAAGCTTTTCGCTTGGAGTT 846
QY 799 ACTATTCTTCTATCGCTGACGGTGTTCCTCAACCTGCTAGCCGAGACCTTGCACACAGTC 858
Db
QY 847 ACAATTCATTATCGCTTACAGTCTTCTCAACATGCTGCGCGAACAATGCCGCGGACC 906
QY 859 TCGACGCTATCCCTCTGTAGGAGCTA 887
Db
QY 907 TCGATGCGGTACCGCTGCTCGGTAAGTA 935

RESULT 5
AAZ24475
ID AAZ24475 standard; cDNA to mRNA; 2886 BP.
XX AC
XX AAZ24475;
XX
XX DT 17-FEB-2000 (first entry)
XX
XX D. melanogaster acetyl-choline receptor DNA from clone Da7.
XX
XX Acetyl-choline receptor; nicotinic; insect; insecticide; screening;
XX neurotransmission; plant protection agent; conductance; AChR; ds.
XX
XX Drosophila melanogaster.
XX
XX Key Location/Qualifiers
XX CDS 372..2684
XX /*tag= a
XX /*product= "acetyl choline receptor."
XX
XX DE19819829-A1.
XX
XX PD 11-NOV-1999.
XX
XX PF 04-MAY-1998; 98DE-01019829.
XX
XX PR 04-MAY-1998; 98DE-01019829.
XX
XX (FARB) BAYER AG.
XX
XX Adamczewski M, Oellers N, Schulte T;
XX WPI; 2000-014207/02.
XX P-PSDB; AAY50814.
XX
XX New nucleic acid encoding a nicotinic acetylcholine receptor from
XX insects, used to identify potential insecticides.
XX
XX Claim 1a; Page 8-12; 26pp; German.
XX
XX This invention describes a novel nucleic acid (NA) encoding a nicotinic
XX acetyl-choline receptor (I) from insects which can be used as an
XX insecticide. Inhibitors of (I) interfere with neurotransmission. (I)

(also vectors containing it, its regulatory regions, and antibodies directed against (i)-encoded proteins) are used to screen for: (a) plant protection agents that alter conductance of AChR, potentially useful as insecticides, or (b) genes which encode polypeptides that are involved in formation of functionally related AChR in insects. (i) are also used to isolate and characterize the specified regulatory regions and for recombinant production of (ii). This sequence encodes an acetyl-choline receptor isolated from *Drosophila melanogaster*

Sequence 2886 BP; 942 A; 687 C; 640 G; 617 T; 0 U; 0 Other;

Query Match 27.4%; Score 411.4; DB 3; Length 2886;
 Best Local Similarity 68.8%; Pred. No. 8.8e-104;
 Matches 565; Conservative 0; Mismatches 256; Indels 0; Gaps 0;

QY 57 AGGTCCTCAGAGAGAGATCTCGAAGCGGTGCTGGCGAAGTACACACCCCTGGAGCG 116
 DB 1295 AGGATATCATGAAGAGAGCTGTACACGATCTTTTGGATCCTTATATACACTAGAACG 1354
 QY 117 ACCGGTGGCCAAACGAGAGCGAACCGCTAGAGGTGAGGTTCGGCTTGACCTTGACGCAAAAT 176
 DB 1355 TCCGGTTCTCAATGAATCGGACCGCTTACAAATTAAGCTTTGGTTAACTTTAATGCAAT 1414
 QY 177 CATTCAGTGGACGAGAGAAATCAACTTATTAACCAATATATGGCTGCTGTTGGAGTG 236
 DB 1415 TATCGATGTGGACGAGAGAAATCAATTCGTAGTCACTAATGTGTGTTAAACTGGAGTG 1474
 QY 237 GAATGACTACACTGAGGTGGACGACGAGGATGCGGGGTCAAGACCTCAGGAT 296
 DB 1475 GAACGACATGAATCTCGCTGGAAACACTCCGACTATGGCGGAGTTAAGGATCTGCCAAT 1534
 QY 297 CAGGCCCAACAAAGTTGTGGAAGCGGACGCTCTTATGTATTAATAGTGTGACAGGAGTTT 356
 DB 1535 ACCGCCCATCGCATCTGGAAGCGGACGCTGCTGATGACAGTGGGATGAGGAT 1594
 QY 357 TGACGGACCTACAGACCAACGTTGGTGGTTCAGAAAGCGGGCGAGTTCCTGTACGTGCC 416
 DB 1595 TGACGGACCTACAGACCAACGTTGGTGGTTCAGAAAGCGGGCGAGTTCCTGTACGTGCC 1654
 QY 417 ACCTGGCATATTAAGAGGACATCAAGATGGACATCGCTGGTTCCTTCGACGACCA 476
 DB 1655 GCGGGGATCTTCAAGTCGACGTCGAGATGCAAGATCGATCACTGCTGCTTCGATGACCA 1714
 QY 477 AACTGTGATGAAGTTCGTAGCTGGACATATGACGGAATCAGTTGATCTGTGCT 536
 DB 1715 GCGGTGGAGATGAAGTTCGACGTTGACCTACGAGGATTCAGCTGGATTTACAAT 1774
 QY 537 AAAGATGAGGACGGCGGATCTATCGGACTTATCAAAATGGGAGTGGTATCTAAT 596
 DB 1775 ACAAGATGAACCTGGCGGTGATATCAGCAGTTACGTCTCAACGGCGAGTGGGAACCTACT 1834
 QY 597 AGGAATCCAGGCAAAAGAACCAATAACATACGCTGCTGCCCCGAGCCCTACGTGGA 656
 DB 1835 GGGTGTCCCGGAAAGCTAACGAGATCTATTACACTGCTGCGGAAACCTATATAGA 1894
 QY 657 CGTCACCTTCACCATCATGATTAAGAGACGAACCTTGTACTTCTTCAACCTGATCGT 716
 DB 1895 CATCACCTTCGCCATCATATCCGCCGACGAACACTGTACTTATTTTCAACCTGATCAT 1954
 QY 717 CCGTGGTGTGATCTATCATGATGGCACTCTCTGGCTTACATCGCACGACGATCCGG 776
 DB 1955 ACCTTGTGTATGATTCCTCCATGCGCTGCTGCGGATTCACCTCGCGCAGATTCGGG 2014
 QY 777 AGAGAACTCACCTTGAGTGCATATTCTTCTATCGTCAAGCTGTTCCTCAACCTGCT 836
 DB 2015 TGAATAATTATCGTGGGTGTTACATCTTGCTCTCGTCAAGCTGTTCTGAATATGTT 2074
 QY 837 AGCCGAGACCTTCGACAGGTCCTCCGACGCTATCCCTCGT 877
 DB 2075 TGCCGAGACATGCGGGCTACTTCCGATGCGGTGCCATTGT 2115

AAC58395
 ID AAC58395 standard; cDNA; 1509 BP.
 XX AAC58395;
 AC AAC58395;
 DT 29-JAN-2001 (first entry)
 XX Human PRO2145 nucleotide sequence SEQ ID NO:76.
 XX
 KW Human; tumour; diagnosis; neoplastic disease; neoplastic cell growth;
 KW proliferation; tumorigenesis; identification; cancer; cytostatic;
 KW neurotropic; neuroprotective; antiinflammatory; immunosuppressive;
 KW immunotumulant; antiangiogenic; leukaemia; lymphoid malignancy;
 KW neuronal disorder; glial disorder; astrocytal disorder; angiogenic;
 KW hypothalamic disorder; glandular disorder; macrophagal disorder;
 KW epithelial disorder; stromal disorder; blastocoelec disorder;
 KW inflammatory disorder; immunologic disorder; ss.
 XX
 OS Homo sapiens.
 XX WO200053755-A2.
 XX
 PD 14-SEP-2000.
 XX
 XX 06-JAN-2000; 2000WO-US000376.
 XX
 XX 08-MAR-1999; 99WO-US0005028.
 XX 02-JUN-1999; 99WO-US012252.
 XX 23-JUN-1999; 99US-0141037P.
 XX 07-JUL-1999; 99US-0143048P.
 XX 26-JUL-1999; 99US-0145698P.
 XX 30-NOV-1999; 99WO-US028313.
 XX 05-DEC-1999; 99WO-US030911.
 XX 05-JAN-2000; 2000WO-US000219.
 XX
 XX (GETH) GENENTECH INC.
 XX Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hillan KJ, Roy MA;
 PI Watanabe CK, Wood WI;
 XX
 XX WPI; 2000-572270/53.
 DR P-PSDB; AAB24088.
 XX
 PT Thirty PRO polynucleotides encoding PRO polypeptides, useful in the
 PT treatment, diagnosis and prevention of cancer.
 XX
 PS Claim 50; Fig 57; 286pp; English.
 XX
 CC The present invention describes an isolated antibody that binds to one of
 CC the human PRO proteins designated PRO212, PRO290, PRO341, PRO535, PRO619,
 CC PRO717, PRO809, PRO830, PRO848, PRO943, PRO1005, PRO1009, PRO1025,
 CC PRO1030, PRO1097, PRO1107, PRO1111, PRO1153, PRO1182, PRO1184, PRO1187,
 CC PRO1281, PRO223, PRO339, PRO834, PRO1317, PRO1710, PRO2094, PRO2145 OR
 CC PRO2198. PRO antagonists can be used to inhibit tumour cell growth. The
 CC PRO polypeptides and nucleotides are useful in the treatment, diagnosis
 CC and prevention of cancer. The antibodies and other anti-tumour compounds
 CC maybe used to treat various conditions, including those characterised by
 CC overexpression and/or activation of the amplified PRO genes. Exemplary
 CC conditions or disorders to be treated with such antibodies and other
 CC compounds include benign or malignant tumours (e.g., renal, liver,
 CC kidney, bladder, breast, gastric, ovarian, colorectal, prostate,
 CC pancreatic, lung, vulva, thyroid, hepatic carcinomas, sarcomas,
 CC glioblastomas, and various head and neck tumours), leukaemias and
 CC lymphoid malignancies, other disorders such as neuronal, glial,
 CC astrocytal, hypothalamic and other glandular, macrophagal, epithelial,
 CC stromal and blastocoelec disorders, and inflammatory, angiogenic and
 CC immunologic disorders. AAC5842 to AAC58366 represent PCR primers and
 CC hybridisation probes used in the isolation of the human PRO sequences.
 CC AAC58367 to AAC58396 and AAB24057 to AAB24089 represent human PRO
 CC polynucleotide and protein sequences given in the exemplification of the
 CC present invention
 XX
 SQ Sequence 1509 BP; 297 A; 451 C; 429 G; 332 T; 0 U; 0 Other;

Query Match		25.1%;	Score 377.2;	DB 3;	Length 1509;
Best Local Similarity		55.2%;	Pred. No. 2.2e-94;		
Matches 811;		Conservative	0;	Mismatches 633;	Indels 24; Gaps 3;
QY	33	GGCTTTGCTGCCCCGTATCGAGCAAGTCTCTACAGAGAGAGACTCCTGAACCGCTTGCT	92		
Db	42	GTGCTCTCTGCAGTGTCTCCCTGCAAGCGGAGTTCAGAGGAGCTTTACNAGAGCTGGT	101		
QY	93	GGCGAACTACAACCCCTGAGGCGACCGGTGGCCAAACGAGAGCGAACCGCTAGAGGTGAG	152		
Db	102	CAAGAACTACAATCCCTTGGAGAGGCCCGGTGGCCAAATGACTCTGCAACCACTCAACCGTCTA	161		
QY	153	GTTCGGCTGACCTTCGAGCAATCATAGCGTGGAGCGAGAACTCACTATTATAC	212		
Db	162	CTTCTCCCTGAGCCTCTCGAGATCATGGACGTGGATGAGAACCAAGTTTAAACCAC	221		
QY	213	CAATATATGCTGCTGTGGAGTGGATGACTACAACCTGAGGTGAAACGACAGCGAGTA	272		
Db	222	CAACATTTGGCTCAAAATGCTTTGGACAGATCACTATTTACAGTGGAAATGTGTCAATA	281		
QY	273	TGGCGGGGTCAAGGACTCTAGATACAGCCCAACAAAGTTGTGGAAGCCGGAAGTCTTTAT	332		
Db	282	TCAGGGGTGAGACTGTTCTGTTTCCAGATGGCCAGATTGGAAACCAAGACATTTCTTCT	341		
QY	333	GTATAATAGTCTGACGAGGTTTTCACGGGACCTACCAGCAACAGTGTGTCTAGAG	392		
Db	342	CTATAACAGTGTGATGAGCGCTTTGACGCCCAITTCACACTAAAGTGTGTGTGAATTC	401		
QY	393	CGCGGCGAGTTGCTCTAGTGCACCTGCGCATATTCAGAGCAGACATGCAAGATGGACAT	452		
Db	402	TTCTGGGCATTGGCAGTACTGCTCCAGGCATATTCAGAGTTCCTGCTACATCATGT	461		
QY	453	CGGTGGTTCCCTTCGAGCAACAACTGTGTATGAAAGTTTCGGTAGCTGGACATATGA	512		
Db	462	ACGCTGGTTTCCCTTTGATGTGAGCACTGCAAACTGAAGTTTGGGTCCTGGTCTTACGG	521		
QY	513	CGSCAATCAGTTGGATCTGGTGTAAAGATAGGCGAGCGGCGATCTATCGACATTCAT	572		
Db	522	AGGCTGGTCTTGGATCTGAGATGAGGA-----GGCAGATATCAGTGGCTATAT	572		
QY	573	AACAAATGGGAGTGTATCTATAGGAATGCCAGGCAAAAGAACACAAATACATACGC	632		
Db	573	CCCCAATGAGAAATGGACCTAGTGGAAATCCCCGGCAAGAGAGTGAAGTTCTATGA	632		
QY	633	GTGCTCCCGGAGCCCTAGTGAAGCTCACCTTCACCATCATGATGAAGAGCAACCTT	692		
Db	633	GTGCTGCAAGAGCCCTACCCCGATGTCACTTTCAAGTGAACCATGCGCGCAGGACGCT	692		
QY	693	GTACTACTTCTTCAACCTGATGCTCCGTGCTGCTGATCTCATCGATGGCACTCTCGG	752		
Db	693	CTACTATGGCTCAACTGCTGATCCCTGTGTGTCTCATCTCCGCCCTCGCCCTGCTGGT	752		
QY	753	CTTCACACTGCCACCACTCCGAGAGAACTCACACTTTGGAGTCACTATTCTTCTATC	812		
Db	753	GTTCCTGCTTCTGCAAGTTCGSGGAGAGATTTCCTCGGGATAACAGTCTTACTCTC	812		
QY	813	GCTGACGGTGTCTCTCAACCTGTGTAGCGAGCCCTGGCCACAGGTCTTCGACGCTATCCC	872		
Db	813	TCTTACCGTCTTCAATGCTGCTGTGTGTGAGATCATGCCCGCAACATCCGATTCGGTACC	872		
QY	873	CTGTTTAGGAGCTTCTCAATTTGCATCATGTTTCATGGTAGCGTCTCTGTGTACTGAC	932		
Db	873	ATTGATAGCCAGTACTTTCGCAAGCACATGATCATCTGTGGGCCCTCTCGTGTGTGTGAC	932		
QY	933	TGTGTGGTACTCAATTAACACCATCGAACAGCTGATATACATGAATGCGACAGTGGAT	992		
Db	933	GGTATCGTGTCTCAGTATACACCAACACAGCCCGAGCGGGGCAAGTGCCTCAAGTGGAC	992		
QY	993	AAAATCAGTATTCCTACATGTTGCCATGGAATCTCGGAAATCTCGAGGCCAGGGAAGAA	1052		
Db	993	CAGAGTCATCTTCTGAACTGGTGCAGTGTGTTTCTCGCAATGAAGAGGCCCGGGGAGGA	1052		

QY	1053	GATCACCCAGAGACTATAATGATGAACACAGGAGATGAGGAGCTGGAACCTGAGGAGAG	1112		
Db	1053	CAAGTGCCTCCCGCTGCTCCAGCAACAGCAGCGCGCTGTCAGCTGGCCAGTGTGGAGAT	1112		
QY	1113	GTGCTCGAAGTCTTCTGTCGCGAATGTTCTAGATATTGATGATCTTCAAGACACGCGCC	1172		
Db	1113	GAGCGCGTGGCGCGCCGCCAGCAACGGGAACCTGCTGTATCATCGGCTTCCGCGG	1172		
QY	1173	TCCGCTCTTAACAGTACTGCTCGACCGGGAATTTGGGACCTGGGTCTCAATATTCG	1232		
Db	1173	CCTGGACGGGTGCACTGTGTCCGACCCCGCACTCTGGGGTAGTGTGTGGCGGCAT--G	1230		
QY	1233	CAGGATTTCCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1292		
Db	1231	GCTGCTCCCCCACGACGATGAGCACCTCTGACGCTGGGCAACCCCCGAGGGGG--	1288		
QY	1293	GGGTAGCCACCATCGAGCTGCACCTCATCTACTGAGAGAGCTGCAGTTTCATCACGCGCAG	1352		
Db	1289	-----ACCCGACTTGGCCAAAGATCTGAGGAGGTCCGCTACATTGCCAACC	1337		
QY	1353	GATGAAGAAGGTGATGAGGAAGCCGAGCTGATCAGCACTGGAAGTTTGTCTCGATGTT	1412		
Db	1338	CTTCCGCTGCCAGACGAAAGCGCGGTCTGACGAGTGGAGTTCGCGCCTGTGT	1397		
QY	1413	TGTTGATAGTGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1472		
Db	1398	GGTGGACCCCTGTGCTCTCATGCTTCTGCTGCTTTCACAACTCATCGACAGTACTGT	1457		
QY	1473	CTGTTATCGGACCCCATATCATCTGTG	1500		
Db	1458	CCTGATGTCGGCTCCCAACTTCTGTGAG	1485		

RESULT 7
AAC90380
ID AAC90380 standard; cDNA; 1509 BP.

XX AAC90380;
XX
XX 14-MAR-2001 (first entry)
XX
XX Wild-type human alpha7 ligand gated ion channel coding sequence.
XX
XX Human; alpha7 nicotinic acetylcholine gated ion channel;
XX 5-hydroxytryptamine; 5-HT3; calcium ion conductance; ss.
XX Homo sapiens.
XX WO200073431-A2.
XX
XX 07-DEC-2000.
XX
XX 25-MAY-2000; 2000WO-US011862.
XX
XX 27-MAY-1999; 99US-0136174P.
XX
XX (PHAA) PHARMACIA & UPJOHN CO.
XX
XX Groppi VE, Wolfe ML, Berkenpas MB;
XX
XX WPI; 2001-061524/07.
XX P-PSDB; AAB50012.
XX
XX Special cell culture medium for treating cells and for inducing mammalian
XX cell lines to conduct calcium ions, comprising specified concentrations
XX of ions of sodium, calcium and potassium at specified pH.
XX
XX Example 5; Page 60-61; 77pp; English.

XX The present sequence is the coding sequence for wild-type human alpha7
XX nicotinic acetylcholine gated ion channel. The human alpha7 ion channel
XX was used in the construction of an alpha7/5-hydroxytryptamine (5-HT3)
XX chimeric ligand gated ion channel (see AAC90382 and AAB50014). The

XX 07-JUN-1996; 96US-00660451.
 XX 07-JUN-1995; 95US-00484722.
 XX (MERI) MERCK & CO INC.
 XX Elliott KJ, Harpold MM;
 XX WPI; 2003-511917/48.
 XX P-PSDB; ADA10874.
 XX
 XX New nucleic acid encoding an alpha-6 or a beta-3 subunit of a human
 XX neuronal nicotinic acetylcholine receptor (nAChR), useful for identifying
 XX compounds that modulate human neuronal nAChR activity.
 XX
 XX Claim 33; Col 67-72; 63pp; English.
 XX
 XX The invention relates to an isolated nucleic acid molecule comprising a
 XX nucleotide sequence encoding an alpha 6 or a beta 3 subunit of a human
 XX neuronal nicotinic acetylcholine receptor nAChR. nAChR's form ligand-
 XX gated ion channels that mediate synaptic transmissions between nerve and
 XX muscle and between neurons upon interaction with the neurotransmitter
 XX acetylcholine. The nucleic acid molecule is useful for identifying
 XX compounds that modulate human neuronal nAChR. The present sequence
 XX represents DNA encoding the human neuronal nicotinic acetylcholine
 XX receptor, nAChR, alpha 7 subunit. Note: the present sequence decodes to
 XX the protein shown in ADA10874 not the one described in the specification
 XX as being the nAChR alpha 7 subunit ADA10865.
 XX
 XX Sequence 1876 BP; 369 A; 553 C; 531 G; 423 T; 0 U; 0 Other;
 XX
 XX Query Match 24.9%; Score 374; DB 8; Length 1876;
 XX Best Local Similarity 55.1%; Pred. No. 1.9e-93;
 XX Matches 809; Conservative 0; Mismatches 635; Indels 24; Gaps 3;
 XX
 XX 33 GCGTTTCTGCTCGGATATCGGAGCAAGTCTCTACGAGAGAGACTCTCTGAACGCGTTGCT 92
 XX 114 GTCGCTCTGACGCTGTCCTGCAAGCGGAGTCTCCAGAGGAAGCTTTTACAGGAGCTGGT 173
 XX 93 GCGCACTCAACACCTCGAGCGGCGGTTGGCCACGAGAGCGAAGCGGTAGAGTCAAG 152
 XX 174 CAGAACTCAATCCCTTGAGAGGCGGCGGTTGGCCAAATGATCGCAACCACTACCGGCTA 233
 XX 153 GTTCGGCTTGACCTTGACGAAATCAATTGACGTGGAGGAGCAAGAAATCAATCTATAAC 212
 XX 234 CTTCCTCTGAGCTCTCTGAGATCATGAGCTGAGTGGATGAGAGCAAGCTTTTAAACAC 293
 XX 213 CAATATATGCTGTGTTGAGTGGAAATGACTCAACCTGAGTGGAAACGACAGCGAGTA 272
 XX 294 CAACATTTGGCTGCAAAATGCTTTGGACAGATCACTATTATACAGTGGAAATGTGCAGAA 353
 XX 273 TGGCGGGTCAAGGACTCAGGATCAGCCCAACAGTGTGGAAGCGGAGCGTCTTAT 332
 XX 354 TCCAGGGTGAAGCTGTGCTGTTTCCAGATGGCCAGATTGGAACCAACAGACTTCTCT 413
 XX 333 GTATAATAGTGTGACGAGGTTTTCAGCGGAGCTTACAGACCAACCTGGTGTGTCAGAA 392
 XX 414 CTATAACAGTGTGATGAGCGCTTTGAGCGCCACATTCACACTAACCTGTTGGTGAATTC 473
 XX 393 CGCGCGGAGTGTGCTGTAGTGGCACTGCGCATATTCAGAGCAATGCAAGATGACAT 452
 XX 474 TTCTGGGCAATTCGACGTACCTGCTCCAGGCATATTCAGAGATTCTCTGTACATCGATGT 533
 XX 453 CGCGTGTGTTCCCTTCGACCAACCACTGTGATGTAAGTTCGGTAGCTGGACATATGA 512
 XX 534 ACCTGTTTCCCTTTGATGTGAGCACTGCAACTGAACTTTGGTCTCTGGTCTTACCG 593
 XX 513 CGGCAATCAGTGTGATCTGTGTTAAAGATAGGCGGCGGCTATCTATCGGACTTCAT 572
 XX 594 AGGCTGTCTCTGATCTGAGATGCGAGGA-----GGCAGATATCAGTGGCTATAT 644
 XX 573 AACAAATGGGAGTGGTATCTAATAGATGCCAGGCAAAAGAACAAATATACATACGC 632

Db 645 CCCAATGGAGATGGGACCTAGTGGGAATCCCGGCAAGAGGAGTGAAGGCTTCTATGA 704
 Qy 633 GTGCTGCCCGAGCCCTAAGTGGAGCTCACTTCAACCATATGATATAAGAACGACCTT 692
 Db 705 GTGCTGCAAGAGAGCCCTACCCCGATGTCACTTCAAGTACCACATGCGCCGACGCT 764
 Qy 693 GTACTACTTCTCAACCTGATCGCTCGGTGGTGTGATCTCATCTGATGGGACTCTCTCGG 752
 Db 765 CTACTATGGCTCTCAACCTGCTGATCCCTGTGTGCTCATCTCCGCTCTGCGCTGTGCT 824
 Qy 753 CTTTACACTGCTCCAGACTCCGAGAGAACTCACACTTGGAGTCACTATTCTTCTATC 812
 Db 825 GTTCTCTCTTCTGAGATTCGCGGAGAAATTTCCCTGGGATACAGCTTCTACTCTC 884
 Qy 813 GCTGACGGTGTCTCAACCTGGTAGCCGAGACCTCTGCCACAGCTCTCCGACGCTATGCC 872
 Db 885 TCTTACCGTCTTCATGCTGCTGCTGATCATGCTCCGCAACATCCGATTCGGTACC 944
 Qy 873 CTTGTTAGGACGCTACTTCAATTCATGTCATGTTTCATGCTAGCGTCTGTGGTACTGAC 932
 Db 945 ATTGATAGCCGAGTACTTCCGAGCACCATGATCATCTGGGCTCTCGGTGGTGGTAC 1004
 Qy 933 TGTGTTGGTACTCAATTTACCACTCGAACAGCTGATATATATGATGATGATGATGATGAT 992
 Db 1005 GGTGATCGTCTGAGTACCAACCAACGACCCGCGGCGCAAGATGCCAAGTGGAC 1064
 Qy 993 AATATCAGTATCTTCTACAAATGGTTCATGATGATGATGATGATGATGATGATGATGAT 1052
 Db 1065 CAGAGTATCTCTTCTGAACTGGTGGCGGTGTTCTCTGGAATGAAGAGGCGCGGGAGGA 1124
 Qy 1053 GATCACCGAGAGACTATATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1112
 Db 1125 CAGGTGCGCGCGCTCTGCGAGCAAGAGCGCGCTGCGAGCTGCGCAGTGTGGAGAT 1184
 Qy 1113 GTCGTGGAAGTCTCTGCTGGCAATGTTCTAGATATTTGATGATGATGATGATGATGATGAT 1172
 Db 1185 GAGCGCGTGGCG 1244
 Qy 1173 TCGCGCTCTCAACAGTACTGCTGCGTGGACCGGGAATTTGGGACCTGGGTGCTCAATATTC 1232
 Db 1245 CCTGGAGCGGCTGCACTGTGTCCGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1302
 Qy 1233 CACGATTTTCCGCTCGGTGCTGCTGCGTCCGTCACGATGGAAGAGCTGGCGCGCGGCT 1292
 Db 1303 GCCTGTCTCCCGACGACGATGAGCACCTCTGCGACGCGCGGCAACCCCGCGAGGGGG-- 1360
 Qy 1293 GGGTAGCCCACTCGGAGCTGCACCTCATATCTAGAGAGCTGCACTTCAATCAAGCGGCGAG 1352
 Db 1361 -----ACCGGACTTGGCCAAAGATCTCTGGAGAGGTCCGCTACATTTGCCAATCG 1409
 Qy 1353 GATGAAGAAGCTGATGAGGAACCGAGCTGATCAGGCACTGGAGTTTCTGCGGATGGT 1412
 Db 1410 CTTCCGCTGCCAGGACGAAAGCGAGCGGTCTCGACGAGTGGAAAGTTCGCCGCTGT 1469
 Qy 1413 TGTGTATAGTGTGCTGCTGTTTCTGTTTCACTTTTCACTTTTCACTTTTCACTTTTCACTTT 1472
 Db 1470 GGTGGACCGCTGTGCTCATGCGCTTCTCGGTCTTCCAGCATCATCTGACCATCATCTGAC 1529
 Qy 1473 CTTGTTATCGGACCGCATATCATCTGTG 1500
 Db 1530 CTTGATGTGGCTCCCAACTTCTGTGGAG 1557

RESULT 10

AAV44687

ID AAV44687 standard; cDNA; 1590 BP.

XX

AAV44687;

XX

DT 09-OCT-1998 (first entry)

XX

DE V274T variant human alpha7 nAChR coding sequence.

Db 1348 TCGCTGCCAGGACGAAGCGGCGTCTGCAGCGAGTGGAAGTTCGCGCGCTGTGTGG 1407
QY 1415 TTGATAGTGTTCCTGCTGTTGTTACACTTTTCAATCATCGGACAGTAGCTGCC 1474
Db 1408 TGGACCGCTGTGCTCATGCGCTTCCTCGTCTTTCACCATCATCTGCACCATCGGCATCC 1467
QY 1475 TGTATCGGCACGCATATCATGCTG 1500
Db 1468 TGAATCGGCTCCCACTTCGTGGAG 1493

RESULT 11

AAV12197
ID AAV12197 standard; cDNA; 1876 BP.

XX
AC AAV12197;

XX 14-MAY-1998 (first entry)

XX Human neuronal nicotinic acetylcholine receptor alpha-7 subunit cDNA.

XX Human; neuronal nicotinic acetylcholine receptor; alpha-7 subunit;

XX brain tissue; screening; NACHR; antibody; ds.

XX Homo sapiens.

XX Key Location/Qualifiers
XX CDS 73..1581

FT /*tag= a

FT /product= "neuronal nicotinic acetylcholine receptor

FT alpha-7 subunit"

XX WO9420617-A2.

XX 15-SEP-1994.

XX 08-MAR-1994; 94WO-US002447.

XX 08-MAR-1993; 93US-00028031.

XX (SALK) SALK INST BIOTECHNOLOGY IND ASSOC.

XX Elliott KJ, Ellis SB, Harpold MM;

XX WPI; 1994-303024/37.

XX P-PSDB; AAW44153.

XX Human neuronal nicotinic acetylcholine receptor subunits and DNA - also
PT transformed cells useful for screening cpds. which modulate activity of
PT the receptor.

XX Claim 8; Page 78-79; 99pp; English.

XX The present sequence encodes a human neuronal nicotinic acetylcholine
CC receptor (NACHR) subunit. The cells expressing the alpha and/or beta
CC NACHR subunits may be used in a method of screening compounds to identify
CC any which modulate the activity of human neuronal NACHR. Subunit specific
CC antibodies may be used to monitor the distribution and expression density
CC of various subunits in normal vs diseased brain tissues. Testing of
CC single receptor subunits or specific receptor subunit combinations with a
CC variety of potential agonists or antagonists provides information with
CC respect to the function and activity of the individual subunits and
CC should lead to the identification and design of compounds that are
CC capable of very specific interaction with one or more receptor subtypes.
CC The resulting drugs should exhibit fewer unwanted side effects than drugs
CC identified e.g. screening with cells that express a variety of subtypes

XX SQ Sequence 1876 BP; 369 A; 553 C; 530 G; 423 T; 0 U; 1 Other;

Query Match

Best Local Similarity 55.08; Pred. No. 2.5e-93;

Matches 808; Conservative 1; Mismatches 635; Indels 24; Gaps 3;

QY 33 GGCTTTGCTGCGGTATCGGAGCAAGGTCTCTACGAGAAGAGACTCTCTGACGCGTTGCT 92
Db 114 GTGCTCTCTGCAAGCTGTCCCTGCAAGCGAGTTCAGAGAAAGCTTTTCAAGAGGACTGGT 173
QY 93 GGGAACTCAACACACCCCTGGAGCGACCGGTGGCCCAACGAGAGCGAACCGCTAGAGGTGAG 152
Db 174 CAAGAACTCAATCCCTTTGGAGAGGCCGCTGGCCCAATGACTCGCAACCACTCACCGTCTA 233
QY 153 GTTCGGCTTGACCTTGAGCAAAATCATTTGACGTGGGACGAGAAGAAATCAACTTATTAAAC 212
Db 234 CTTCTCCCTGAGCCTCTCTGAGATCATGACGCTGGATGAGAAGAACCAAGTTTTTAAACAC 293
QY 213 CAATATATGCTGCTGTTGGAGTGAATGACTACAACCTGAGTGGACACAGCGAGTA 272
Db 294 CAATTTGGCTGCAAAATGCTTGGACAGATCACTATTTACAGTGAATGTGTGAGATA 353
QY 273 TGGCGGGTCAAGGACCTCAGGATCACGCCCAACAAAGTTTGGAAAGCCGACCTCTTAT 332
Db 354 TCCAGGGTGAAGACTGCTGTTTCCAGATGCCAGATTTGGAAACCAAGACATCTTCT 413
QY 333 GTATATAGTGTGACGAGGTTTGTGACGGACCTTACAGACCACTGAGTGTGTCAGAAG 392
Db 414 CTATAACAGTGTGATGAGCGCTTTGACGCCACATTCACACTAACCTGTTGGTGAATTC 473
QY 393 CGGCGCAGTTGCTGTACGTGCCACCTGTCATATTCAGAGACACATGCAAGATGGACAT 452
Db 474 TTCTGGCATTTGCCAGTACTGCTCTCAGGCAATATTCAGAGTTCCTGTCTACATCGATGT 533
QY 453 CGCGTGTTTTCCCTTCGACGACCAACACTGTGATATGAAGTTGCGTGTGAGTGGACATAGA 512
Db 534 ACGCTGTTTTCCCTTTGATGTGCAGCACTGCAAACTGAAGTTTGGTGGTCTTACGG 593
QY 513 CGGCAATCAGTTGGATCTGCTGTAAAGATGAGGAGCGGCGGATCTATCGGACTTCAT 572
Db 594 AGGCTGCTCTTGATCTGCAGATGCAGGA-----GCGAGATATCAGTGGCTATAT 644
QY 573 AACAAATGGGGAGTGTATCTAATAGGAATGCCAGGCAAAAAGAAACAATAACATACGC 632
Db 645 CCCNATGGAGATGGACCTAGTGGAAATCCCGGCAAGAGGATGAAGTTCATATGA 704
QY 633 GTGCTGCCCGAGCCCTTACGTGGACGTCACCTTTCACCATCATGATATAAGAGACGAACCTT 692
Db 705 GTGCTGCAAGAGAGCCCTTACCCCGATGTCACTTTCACAGTGCACCATGCGCGCAGACGCT 764
QY 693 GTACTACTTCTTCAACCTGATCTCCGTCGGTGTCTGATCTCATCGATGCGACTCTCTCGG 752
Db 765 CTACTATGGCTCAACCTGCTGATCTCCCTGTGTGCTCATCTCCGCCCTCCGCTGCTGGT 824
QY 753 CTTTCACACTGCCACCACTCCGAGAGAAACTCACACTTTGGAGTCACTATTCTTCTATC 812
Db 825 GTTCCTGCTCTTCGAGATTCGGGGAGAGATTTCCCTGGGGATACAGTCTTACTCTC 884
QY 813 GCTGACGGTGTCTCTCAACCTGGTAGCCGAGACCCCTGCCACAGGTCTCCGACGCTATCCC 872
Db 885 TCTTACGCTCTTCATGCTGCTGCTGGCTGAGATCATGCCGCCAACATCCGATTCGGTACC 944
QY 873 CCGTGTAGGACCTACTTCAATTCATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 932
Db 945 ATTGATAGCCAGTACTTCCGCCAGCACCATGATCATCTGGGGCTCTCCGGTGGTGGTGC 1004
QY 933 TGTGGTGTACTCAATTTACCACCATGCAACAGTGTATATACATGAATGCCACAGTGGAT 992
Db 1005 GGTGATCGTCTGAGTACCAACCAACAGCCCGGCGGGGCAAGATGCCAAGTGAC 1064
QY 993 AAATCAGTATTCCTACATGTTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1052
Db 1065 CAGAGTCACTCTTCTGAACCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1124
QY 1053 GATCACGAGGAGACTATATATGATGAACAGGATGAGGAGTGGAGCTGGAACCTGAGGAGAG 1112
Db 1125 CAAGGTGGCGCCGCTGCCAGCAGCAGCGGCGCTGCAGCTGGCCAGTGTGGAGAT 1184

QY 1113 GTGCTGGAAGTCTTCTGCGGAAATGTTCTAGATATATGATGACTTTCAGACACGGCCC 11172
Db 1185 GAGCGCCGTGGCCGCGCCGCGCCAGCAACGGGAACCTGCTGTATCATCGGCTTCGCGG 1244
QY 1173 TCGCCTCTTAAAGTACTGCTTCGACCGGGAATTTGGGACCTGGGTGCTCAATATTCG 1232
Db 1245 CTTGGACGGGTGCACTGTGTCCGACCCCGACTCTGGGGTAGTGTGTGGCCGCAAT--G 1302
QY 1233 CAGGATTTCCGTGCGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1292
Db 1303 GCCTGCTGCCCAACGACGATGACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1360
QY 1293 GGGTAGCCACCATCGGAGCTGCACCTCATCTACTGAGAGAGCTGCAGTTTCATCGGCCAG 1352
Db 1361 -----ACCGGACTTGGCCCAAGATTCCTGGAGAGGTCCGCTACATTCGCAATCG 1409
QY 1353 GATGAAGAAGCTGATGAGGAAGCCGAGCTGATCAGCACTGGAAGTTTCTGCGATGGT 1412
Db 1410 CTTCCGCTGCAGACGAAAGCAGGCGGTCTGACGGAGTGAAGTTGCGCCGCTGTGT 1469
QY 1413 TGTGTAGATGTTTGGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTGCTGTTG 1472
Db 1470 GGTGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1529
QY 1473 CTTGTTATCGGACCGCATATCATCTGTG 1500
Db 1530 CTTGATGCTGGTCCCAACTTCTGTGGAG 1557

RESULT 12

ID ABS54875
AC ABS54875;
XX 06-DEC-2002 (first entry)
XX Human neuronal nicotinic acetylcholine receptor alpha 7 subunit cDNA.
DE Human; neuronal nicotinic acetylcholine receptor; nNACHr; gene; ss;
KW Human; ion flux; alpha 7 subunit.
KW Homo sapiens.
OS Homo sapiens.
XX Key Location/Qualifiers
FH CDS 73..1581
FT /*tag= a
FT /product= "Human nNACHr alpha 7 subunit"
XX US6440681-B1.
XX 27-AUG-2002.
XX 07-JUN-1995; 95US-00487596.
XX 03-APR-1990; 90US-00504455.
XX 30-NOV-1992; 92US-00938154.
XX 08-MAR-1993; 93US-00028031.
XX 08-NOV-1993; 93US-00149503.
XX (MERI) MERCK & CO INC.
XX Elliott KJ, Ellis SB, Harpold MM;
XX WPI; 2002-711528/77.
XX P-PSDB; ABG70492.
XX Identifying antagonists or agonists of human neuronal nicotinic
PT acetylcholine receptors, by contacting recombinant cells with test
PT compound, and measuring ion flux of cells or binding of compound to
PT nNACHr.
XX Claim 101; Col 57-60; 56pp; English.

XX The invention relates to a method for identifying compounds that are
CC antagonists or agonists of human neuronal nicotinic acetylcholine
CC receptors (nNACHrs), by contacting recombinant cells with a test compound
CC and measuring ion flux, the electrophysiological response of the cells or
CC binding of the test compound to the nNACHr. The recombinant cells are
CC produced by transfection with a nucleic acid encoding at least one human
CC nNACHr (alpha or beta) subunit, such that the cells express an nNACHr
CC comprising one human subunit encoded by the transfected nucleic acid.
CC This sequence represents cDNA encoding the alpha 7 subunit of the human
CC nNACHr polypeptide

XX Sequence 1876 BP; 369 A; 553 C; 530 G; 423 T; 0 U; 1 Other;

Query Match 24.9%; Score 373.6; DB 6; Length 1876;
Best Local Similarity 55.0%; Pred. No. 2.5e-93;
Matches 808; Conservative 1; Mismatches 635; Indels 24; Gaps 3;

QY 33 GGCTTTGCTGCCGTATCGGACCAAGTCTCTCAGAGAGAGACTCTCTGAACGGTTGCT 92
Db 114 GTCGCTCTCTGACGTCTCTCTGAGAGGCGCGTGGCCATGACTCGCAACCACTCCCGTCTA 173
QY 93 GCGCAACTACAACACCCCTGGAGCGACCGGTGGCCAAACGAGAGCGAACCGCTAGAGGTGAG 152
Db 174 CAAGAACTACATCCCTTGGAGAGGCGCGTGGCCATGACTCGCAACCACTCCCGTCTA 233
QY 153 GTTCGGCTTGACCTTGGAGCAATCATTTGACGTGGACGAGAACTCACTACTTATAAC 212
Db 234 CTTCTCCTTGAGCCTCTGACATCATGACGTGGATGAGAGAGCAAGTTTAAACCAC 293
QY 213 CATATATGGCTGCTGTTGGAGTGAATGACTACACCTGAGGTGGAACGACAGCGAGTA 272
Db 294 CAACATTTGGCTGCAAAATGTTTGGACAGATCACTATTTACAGTGAATGTGTGAGAATA 353
QY 273 TGGCGGGGTCAAGGACCTCAGATCACGCCCAAGATTGTGGAAGCGGACGCTCTCTAT 332
Db 354 TCCAGGGGTGAGACTGTTCTGTTCCAGATGCCAGATTGGAACACGACATCTTCTCT 413
QY 333 GTATAATAGTGTGACGAGGGTTTTCACGGGACCTACAGACCAACGTTGTTGTCAGAA 392
Db 414 CTATAACAGTGTCTGATGAGCGCTTTGACGCCACATTCACACATTAAGCTGTTGTTGAATTC 473
QY 393 CGCGGCGAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 452
Db 474 TCTGCGGCAATGCGGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 533
QY 453 CGCGTGGTTTCCCTTCGACGACCAACACTGTGATATGAAGTTTCGGTAGCTGACATATGA 512
Db 534 ACGCTGGTTTCCCTTTGATGTGCGACACTGCAAACTGAAGTTTGGGTCTGCTTACGG 593
QY 513 CGGCAATCAGTTGGATCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 572
Db 594 AGGCTGGTCTCTGGATCTGCAATGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 644
QY 573 AACAAATGGGAGTGTCTATCTAATAGGAATGCCAGGCAAAAGAACACATATACATACGC 632
Db 645 CCCCAATGAGGAATGGGACCTAGTGGAAATCCCCGGAAGAGAGTGAAGAGTTCTATGA 704
QY 633 GTGCTGCGCGGAGCCCTACGTCGACGCTCACCTTCAACATCATATGAAGAGAGCAACCTT 692
Db 705 GTGCTGCAAGAGCCCTACCCCGATGTCACTTTCAGTCAACCTGCGCGCAGGAGCT 764
QY 693 GTACTACTTCTTCAACCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 752
Db 765 CTACTATGGCTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 824
QY 753 CTTTACACTGCCACAGACTCCGGAGAGAACTCACACTTGGAGTCACTATTCTTCTATC 812
Db 825 GTTCTGCTTCTGCGAGATTCGGGGAGAGATTTCCCTGGGATTAACAGTCTTACTCTC 884
QY 813 GCTGACGGTGTCTCTCAACCTGGTAGCGAGACCTCGCCACAGGCTCTCCGACGCTATCCC 872
Db 885 TCTTACCGTCTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 944

645 CCCAATGAGATGGACCTAGTGGGAATCCCGGCAAGAGAGTGAAGGTTCTATGA 704
633 GTCTGCGCGAGCCCTAGTGACGCTACCTTCAACATCATGATGAAGACGAACCTT 692
705 GTCTGCAAGAGCCCTACCCGATGTACCTTACAGTGACCATGCGCGCAGAGCGCT 764
693 GTACTACTTCTTCAACTGATCGTCCGCTGCGTGTGATCATCATGATGACCTCCCTCG 752
765 CTACTATGGCTCAACTGCTGATCTGTGCTCATCTCGCCCTCGCCCTGCTGGT 824
753 CTTTACACTGCCACGAGCTCCGGAGAGAACTCACACTTTGGAGTCACTATTCTTATC 812
825 GTTCTGCTTCTCGAGATTCCGGGAGAGATTTCCTCGGGATTAACAGTCTTACTC 884
813 GCTGACGGTGTCTCTCAACTGATGCGGAGACCTGCGCACAGGTCTCGACGCTATGCC 872
885 TCTTACCGCTTTCATGCTGCTGCTGAGATCATGCGCCCAACATCCGATTCCGTACC 944
873 CTGTTAGGAGCGTACTTCAATTGCAATGTCATGTTTCATGGTAGCGTCTGTGTACTGAC 932
945 ATTGATAGCCAGTACTTCCGACACCAATGATCATCTGCGGCTCTCGGTGTGTGAC 1004
933 TGTGTTGGTACTCAATTACCACTCGAACAGAGTATATACATGAATGCCACAGTGGAT 992
1005 GGTGATCGTCTGACGTACCAACACAGACCGCGGCGGCAAGATGCCAAGTGGAC 1064
993 AATAACAGTATTCTCAATGTTGCCATGGATCTGCGAATGTGAGGCTCAGGAGAGAA 1052
1065 CAGAGTCACTCTTCTGAACGTGTGCGGTGTCTTCTSCGAATGAAGAGGCGCGGGAGGA 1124
1053 GATCACCAGGAAGACTAATGATGAACAGAGAGTGGAGCTGGAATGAAGAGAG 1112
1125 CAAAGTGGCGCGCTGCGCAGACAGAGACGCGCGCTGAGCTTGCCAGTGTGAGAT 1184
1113 GTCTGCAAGTCTTGTGCGCAATGTTCTAGATATTGATGATGATTCAGACACGCGCC 1172
1185 GAGCGCGTGGCGCGCGCGCGCGAGCAACGGAACCTGCTGATCATCGCTTCCGCG 1244
1173 TCGCCTCTTAACAGTACTGCTCGACCGGGAATTTGGGACCTGGGTGCTCAATATTCGG 1232
1245 CTTGACGCGGTGCACTGTGTCTCCGACCGCCCGACTCTGGGGTAGTGTGCGCGCAT -G 1302
1233 CAGGATTTCCGTCGCTGCTGCTCGTCCGTCACGATGGAAGAGTGGCGCGCGGCT 1292
1303 GCTGCTCCCGACGACGATGACACCTCTCGACGCGCGGCAACCCCGCGGGGG-- 1360
1293 GGGTAGCCACATCGGAGCTGCACCTCTACTAGAGAGCTGCAGTTCTATCAGGCCAG 1352
1361 -----ACCGGACTTGGCAAGATCCTGGAGAGGTCCGCTACATTGCCAATCG 1409
1353 GATGAAGAGGCTGATGAGGAACCGAGCTGATCAGGACTGGAAGTTTCTCGGATGGT 1412
1410 CTTTCCGCTGCGAGGACGAAGAGGCGGCTGACGCGAGTGAAGAGTTTCCCGCCTGTGT 1469
1413 TGTGTATAGTTTGTGCTGTGTGTTTACACTTTTACAAATCATCGCAGCAGTAGTGT 1472
1470 GGTGAGCGGCTGTGCTCATGGCTTCTCGGCTTCCCATCATCTGACCAATCGGCAT 1529
1473 CTTGTTATCGGACCGCATATCTGCTG 1500
1530 CTTGATGTGGCTCCCAACTTCGTGGAG 1557

RESULT 14

AAC90385

ID AAC90385 standard; cDNA; 1509 BP.

XX AAC90385;

AC AAC90385;

XX 14-MAR-2001 (first entry)

XX Mutant human alpha7 ligand gated ion channel coding sequence #1.

DE

XX

KW Human: alpha7 nicotinic acetylcholine gated ion channel; mutant;
KW 5-hydroxytryptamine; 5-HT3; calcium ion conductance; ss.
XX Homo sapiens.
OS WO200073431-A2.
PN 07-DEC-2000.
XX 25-MAY-2000; 2000WO-US011862.
XX 27-MAY-1999; 99US-0136174P.
XX (PHAA) PHARMACIA & UPJOHN CO.
XX Groppi VB, Wolfe ML, Berkenpas MB;
PI P-PSDB; AAB50015.
XX WPI; 2001-061524/07.
XX P-PSDB; AAB50015.
XX Special cell culture medium for treating cells and for inducing mammalian
PT cell lines to conduct calcium ions, comprising specified concentrations
PT of ions of sodium, calcium and potassium at specified pH.
XX Claim 50; Page 69; 77pp; English.
XX The present sequence is the coding sequence for a mutant human alpha7
CC nicotinic acetylcholine gated ion channel. The human alpha7 ion channel
CC was used in the construction of an alpha7/5-hydroxytryptamine (5-HT3)
CC chimeric ligand gated ion channel (see AAC90382 and AAB50014). The
CC alpha7/5-HT3 chimeric ion channel can be expressed by recombinant cells
CC in the present invention, resulting in preferential calcium ion
CC conductance by the cells. The protein encoded by this sequence has the
CC wild-type threonine residue at position 230 substituted by a proline
XX residue
SQ Sequence 1509 BP; 296 A; 452 C; 429 G; 332 T; 0 U; 0 Other;

Query Match 24.8%; Score 372.4; DB 4; Length 1509;
Best Local Similarity 55.0%; Pred. No. 4.9e-93;
Matches 808; Conservative 0; Mismatches 636; Indels 24; Gaps 3;

QY 33 GGCTTTGCTGCCGTATCGAGCAAGTCTCTCACGAGAGAGACTCTCTGAACCGTTGCT 92
Db 42 GTCTGCTTCTGACGTGTCCCTGCAAGGCGAGTTCCAGAGGAAGCTTTACAGGAGCTGT 101
QY 93 GGCGAACTACAACACACCTCGAGCGACCGGTGGCCAAACGAGAGCGAACCGCTAGAGTCA 152
Db 102 CAAGAACTACAATCCCTTGGAGAGGCCCGTGGCCAAATGACTCGCAACCACTCACCGTCTA 161
QY 153 GTTCGGCTTGACCTTGCAGCAAAATCATTTGACGTGGACGAGAGAAATCAACTACTTATAAC 212
Db 162 CTTCTCCCTGAGCTCTCTGCAGATCATGACGTGGATGAGAGAAACCAAGTTTAAACCAC 221
QY 213 CAATATATGGCTCTCGTTGGAGTGAATGACTACAACTGAGGTGAACGACGAGGAGTA 272
Db 222 CAACATTTGGCTGCAAAATGTTTGGACAGATCACTATTTACAGTGAATGTGTGAGAAATA 281
QY 273 TGGCGGGGTCAAGGACCTCAGGATCACGCCCAACAAAGTTGTGGAAGCCGAGACGCTCTAT 332
Db 282 TCCAGGGGTGAGACTGTTCTGTTTCCAGATGCCAGATTGGAAACCAAGACATCTCTCT 341
QY 333 GTATAATAGTGTGACGAGGGTTTGAACGGGACCTTACAGACCAACGTCGTTGTGTGAGAAG 392
Db 342 CTATAACAGTGTGTGAGCGCTTTTGACGCCACATTCACACACTTAACGTGTGTGTAATTC 401
QY 393 CGCGCGCAGTTGCTGTAGTGCACCTGCGCATATTTCAAGAGACACATCAAGATGACAT 452
Db 402 TTCTGGGCATTTGCCAGTACTCTGCTCCAGGCATATTTCAAGAGTTCCTGTACATCATGT 461
QY 453 CGCGTGGTTTCCCTTCGACGACCAACACTGTGATATGAAGTTTCGGTAGCTGGGACATATGA 512
Db 462 ACCTGTTTCCCTTTGATGTGAGCAGCTGCAAACTGAAGTTTGGGTCTGCTTACCG 521


```
QY 393 CGGCGCAGTTCCTGTAGTGGCCACTGGCATATTCAAGACACATGCAAGATGACAT 452
Db 402 TTCTGGGCATTGCGAGTACCTGCTCCAGGCATATTCAAGAGTTCTGCTACATCGATGT 461
QY 453 CGCGTGGTTTCCCTTCGAGACCAACACACTGTGATATGAAGTTCCGTTAGTGGACATATGA 512
Db 462 ACCTGGTTTCCCTTTGATGTGAGCACTGCAACTGAACTTTGGGTCTGGTCTTACGG 521
QY 513 CGGCAATAGTTGGATCTGGTGTCTAAAGATGAGGCGGCGGATCTATCGGACTTCAT 572
Db 522 AGCTGGTCTCTGGATCTGAGATGCAGGA-----GGCAGATATCAGTGGCTATAT 572
QY 573 AACAAATGGGAGTGTACTATAGGAATGCCAGGCAAAAGAAACACAATAACATAACGC 632
Db 573 CCCAAATGGGAATGGACCTAGTGGGAATCCCGGCAAGAGAGTGAAGGTTCTATGA 632
QY 633 GTGCTGCCCGAGCCCTAGCTGACGCTCACCTTCACCATCATATAGAAGACGAACCTT 692
Db 633 GTGCTGAAAGAGCCCTACCCCGATGTCACTTCACAGTGACCATGCGCGCAGGACGCT 692
QY 693 GTACTACTTCTTCAACCTGATCTCCGCTGCTGATCTCATCTCATGAGCACTCTCTCGG 752
Db 693 CTACTATGGCCTCAACCTGCTGATCCCGAGTGTCTCATCTCCGCGCTCGCCCTGCTGT 752
QY 753 CTTACACTGCCACAGACTCCGGAGAGAACTCACACTGGAGTCACTATTCTTCTATC 812
Db 753 GTTCTGCTTCTGTCAGATTCCGGGGAGAAATTCCCTGGGGATTAACGTCTTACTCTC 812
QY 813 GCTGACGGTGTCTCTCAACCTGGTAGCCAGACCTCGCCACAGGTCTCCGACGCTATCCC 872
Db 813 TCTTACCGTCTTCATGCTGCTGCTGCTGATCATGCCCCGCAACATCCGATTCCGTTACC 872
QY 873 CCTGTTAGGACGCTACTTCAATTGCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 932
Db 873 ATTGATAGCCAGTACTTCCAGCAGCACCACATGATCATGCTGGGCTCTCTCGTGTGTG 932
QY 933 TGTGTTGTTACTCAATTACCACTCGAACAGCTGATATATACATGAATGCCACAGTGGAT 992
Db 933 GGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 992
QY 993 AAAATCAGTATTCTTCAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1052
Db 993 CAGAGTATCTCTTCTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1052
QY 1053 GATCACCAGGAAGACTATATATGATGAACACAGGATGAGGAGCTGGAATGAAGAGAG 1112
Db 1053 CAAAGTGCGCCCGCTGCGAGCAACAGCAGCGGCTGCGAGCTGCGCAGTGTGGAGAT 1112
QY 1113 GTCGTGGAAGTCTTGTGTCGGAATGTTCTAGATATTGATGACTTCAGACACAGGCC 1172
Db 1113 GAGCGGCTGGCGCGCGCGCGCGCAGCAACGGGAACCTGCTGTACATCGGCTTCGCGG 1172
QY 1173 TCCGCTCTTAAAGTACTGCTTCGACCGGGAATTTGGGACCTTGGGTGCTCAATATTCCG 1232
Db 1173 CCTGGAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1230
QY 1233 CAGGATTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1292
Db 1231 GCCTGCTCCCGCCACGACGATGAGCCTCTGTCAGCGCGGCAACCCCGGAGGGG-- 1288
QY 1293 GGGTAGCCACCATCGCGAGCTGCACCTCATACTGAGAGAGTGCAGTTTCATCAGGCCAG 1352
Db 1289 -----ACCGGACTTTGGCAAGATCTTGGAGAGGTCCGCTACATTCGCAATCG 1337
QY 1353 GATGAGAGAGGCTGATGAGGAAGCCGAGCTGATCAGCGACTGGAAGTTTCTGCGATGGT 1412
Db 1338 CTTCCGCTGCGAGGACGAAAGAGGAGCGGCTGCGAGCGAGTGGAGTTCCGCGCTGTGT 1397
QY 1413 TGTGATAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1472
Db 1398 GGTGAGCCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1457
QY 1473 CCTGTTATCGGACCGCATATCATCTGTG 1500
```

Db 1458 CTTGATGTCGGCTCCCACTTCTGTGGAG 1485

Search completed: May 7, 2004, 15:01:54
Job time : 617.489 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 14:08:32 ; Search time 4039.19 Seconds

(without alignments)

11111.850 Million cell updates/sec

Title: US-09-303-232-5_COPY_95_1597

Perfect score: 1503

Sequence: 1 atggccctatgttggcgcg.....caccgcataatcatcgtgcaa 1503

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Lasting first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estnu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_estl:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_pbg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	423.8	28.2	885	12	BG632919	BG632919 GH16126.3
2	331.2	22.0	607	9	AI292581	AI292581 GH15518.5
3	302.2	20.1	2296	14	CD013901	CD013901 90134548
4	301.2	20.0	1201	9	AL530299	AL530299 AL530299

5	294.8	19.6	1436	29	AY402873	Homo sapi
6	293.2	19.5	2940	11	AK034228	Mus muscu
7	293.2	19.5	3230	11	AK083157	Mus muscu
8	286.8	19.1	1436	29	AY402874	Pan trogl
9	281.6	18.7	3483	11	AK081254	Mus muscu
10	281.6	18.7	4037	11	AK049722	Mus muscu
11	281.6	18.7	4046	11	AK051742	Mus muscu
12	275.6	18.3	1864	11	AK053497	Mus muscu
13	275.6	18.3	2916	11	AK051730	Mus muscu
14	275.6	18.3	3126	11	AK080415	Mus muscu
15	274.4	18.3	1442	29	AY402876	Homo sapi
16	272.4	18.1	1436	29	AY402875	Mus muscu
17	267	17.8	1442	29	AY402877	Pan trogl
18	258.8	17.2	2010	11	AK080475	Mus muscu
19	256.6	17.1	833	14	CB245337	UI-M-FY0-
20	253	16.8	1454	29	AY402878	Mus muscu
21	246.2	16.4	1374	29	AY406232	Mus muscu
22	246.2	16.4	4290	11	AK029177	Mus muscu
23	240	16.0	908	13	BUI49265	AGENCOURT
24	239.8	16.0	1374	29	AY406230	Homo sapi
25	239.2	15.9	922	13	BUI915857	AGENCOURT
26	239.2	15.9	1466	14	CD013891	90139069
27	238.2	15.8	1603	29	AY411327	Homo sapi
28	236.8	15.8	1036	13	BX437801	BX437801
29	234.8	15.6	755	13	BUI702422	UI-M-FY0-
30	231.6	15.4	902	29	AY407186	Mus muscu
31	229.4	15.3	759	14	CD804155	UI-M-GV0-
32	229	15.2	730	14	CF742344	UI-M-HB0-
33	228.2	15.2	615	14	CB149460	K-EST0205
34	228.2	15.2	1374	29	AY406231	Pan trogl
35	220.4	14.7	2513	11	AK033068	Mus muscu
36	219	14.6	1603	29	AY411329	Mus muscu
37	216.6	14.4	902	29	AY407184	Homo sapi
38	216.2	14.4	819	14	CF530720	UI-M-FY0-
39	212.6	14.1	1034	13	BX403124	BX403124
40	210.8	14.0	755	10	AW914206	EST345510
41	210.4	14.0	658	12	BM711715	UI-E-CL1-
42	210.4	14.0	797	14	CA326954	UI-M-FY0-
43	208.8	13.9	777	14	CF538185	UI-M-GI0-
44	205.8	13.7	791	14	CF739825	UI-M-HD0-
45	202.2	13.5	939	13	BQ720344	AGENCOURT

ALIGNMENTS

RESULT 1
BG632919/c
LOCUS
DEFINITION
BG632919 885 bp mRNA linear EST 23-APR-2001
GH16126.3prime GH Drosophila melanogaster head POT2 Drosophila
melanogaster cDNA clone GH16126 3 similar to CG4128: FBan0004128
'ion channel' located on: 2L 30D1-30E1; 04/10/2001, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 885)
Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
Lewis,S. and Rubin,G.M.
BDGP/HMI Drosophila EST Project
Unpublished (2001)
Other ESTs: GH16126.5prime
Contact: Stapleton, M.
BDGP

Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_estfruitfly.berkeley.edu
Based upon the presence of a XhoI site followed by a run of 14 or

more T residues at the beginning of the sequence, this clone was polyadenylated. The resulting Poly-T sequence has been removed. hit genomic AE003511: arm:X (18792641..19136447)
estimated-cyto:18A3-18C6: 04/10/2001
Plate: GH.161 row: C column: 2
High quality sequence stop: 784.

FEATURES

source

1..885 Location/Qualifiers

/organism="Drosophila melanogaster"

/mol_type="mRNA"

/db_xref="taxon:7227"

/clone="GH16126"

/sex="male and female"

/dev_stage="adult"

/lab_host="DH5 - alpha"

/clone_lib="GH Drosophila melanogaster head pot2"

/notes="Organ: head; Vector: pot2; Site 1: EcoRI; Site 2: XhoI; Sized fractionated cDNAs were directly ligated into pot2. Plasmid cDNA library."

ORIGIN

Query Match 28.2%; Score 423.8; DB 12; Length 885;
Best Local Similarity 68.9%; Pred. No. 1.7e-95;
Matches 581; Conservative 0; Mismatches 262; Indels 0; Gaps 0;

QY 161 TGACCTTGACGAAATCATTGACGTGACGAGAGAAATCAACTACTTATAACCAATATAT 220
DB 884 TAACTCATGACGATTATCGATGTGCACGAGAAATCAACTGCTTATACGAATATT 825
QY 221 GGCTGCTGTTGGAGTGGAAATGACTACAACTGAGGTGGACGACGAGTATGGGGGG 280
DB 824 GGCTCAATTGGATGGAACGATATGAATCTTCGATGGAATTCGAGTGAGTTCGGTGTG 765
QY 281 TCAAGGACCTCAGATCAGCCCAACAAGTTGTGGAAGCCGGAAGTCTTATGTATAATA 340
DB 764 TGGGGATCTCGAAATTCGCCACATCGCTATGGAACCGGATGATGATGATCAACA 705
QY 341 GTGCTGACGAGGCTTTGACGGGACCTACCAGACCAACGTTGGTTCAGAACGGGGCA 400
DB 704 GTCCGACGAGGGCTTCGATGGAACGTACCCCAAAATGTTGGTTCGCAATAATGGGA 645
QY 401 GTTGCCTGTACGTGCCACCTGGCATATTCGAAGACACATCAAGATGGACATCGCGTGT 460
DB 644 GCTGTCTGACGTACCGCCAGGTATATTTAAGTCAACGTGTAAGATCGACATACGTGT 585
QY 461 TTCCCTTCGACGACCACTGTGATATGAATTCGTTAGTACGATGACATGACGGCAATC 520
DB 584 TTCCATTCGACGATCAGATGTGAATGAAATTTGGTTTCGTGACCTACGATGGGTTTC 525
QY 521 AGTTGGATCTGTTGCTAAAGATGAGCGCGCGCGATCTATCGGACTTCATTAACAAATG 580
DB 524 AGTTGACCTGAGTTGCAGGACGAAGCTGTTGGCGGACATTTCTAGCTTTATACCAATG 465
QY 581 GGGAGTGGTATCTAATAGGAATGCCAGGCAAAAGAACACAATAACATACGCGTGTGCC 640
DB 464 GCGAATGGGACTTGTAGGTGTGCCCGGTAAACGAAATGAAATCTACTATAATGCTGCC 405
QY 641 CCGAGCCCTACGTGGAGCTACCTTCACCATCATGATGAAGAAGACGACCTTGCTACT 700
DB 404 CAGAACCTTATATGACATACAAATTCGCCATTTTGATAAGCGCAAAACGTTGCTACTAT 345
QY 701 TCTTCAACCTGATCGTCCCGTGGTGTGATCTCATCGATGGGACTCCTCGGTTTCACAC 760
DB 344 TTTTCAATCTGATGTCCCGTGGTGTGATCGCTCCATGACGCTAGGTGTTTACAC 285
QY 761 TGGCACAGACTCGGAGAGAACTCACTTGGAGTCACTATTCTTCTATCGCTGACGG 820
DB 284 TGGCACCAAGATCTGTGTGAAAGCTTTTCGCTTGGAGTTACAATTTCTATTATCGCTTAC 225
QY 821 TGTTCCTCAACCTGAGCGAGACCTGCCACAGGTCTCCGAGCTATCCCTCGTTAG 880
DB 224 TCTTCCTCAACATGGTGGCCGAAACAAATGCGGGGACCTCCGATGCGGTACCGTGTCTG 165

QY 881 GGACGTACTTCAATTGCATCATCTTTCATGTCAGTCGCTGTGTGTGTAAGTGTGGTGG 940
DB 164 GAACATTATTTCATTGATTAATGTTATGTTGTCCTCATCAGTTGTGTCAACCACTATTG 105
QY 941 TACTCAATTACCAATCCATCCAGCTGATATCATGAAATGCCACAGTGGATAAATCAG 1000
DB 104 TCTCAATTATCATCATAGAAATCCAGATACGATGAAATGAGTGAATGGTAAATAGTTT 45

1001 TAT 1003

44 TAT 42

RESULT 2

AI292581

LOCUS

DEFINITION

GH15518.5prime GH Drosophila melanogaster head pot2 Drosophila melanogaster cDNA clone GH15518 5 similar to CG4128: F84004128, 'ion channel', located on: 2L 30D1-30E1; 04/10/2001, mRNA sequence.

ACCESSION

AI292581

VERSION

AI292581.1

KEYWORDS

EST.

SOURCE

ORGANISM

Drosophila melanogaster (fruit fly)

Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

1 (bases 1 to 607)

AUTHORS

Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G., Lewis, S. and Rubin, G.M.

TITLE

BDGP/HMI Drosophila EST Project

JOURNAL

Unpublished (2001)

COMMENT

Other ESTs: GH15518.3prime

Contact: Stapleton, M.

BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: <http://www.fruitfly.org/EST.estimate@fruitfly.berkeley.edu>

hit genomic AE003626: arm:2L [9617316..9882551]

estimated-cyto:30C7-30F4: 04/10/2001

Plate: GH.155 row: B column: 6

High quality sequence stop: 521

POLYA=No.

Location/Qualifiers

1..607

/organism="Drosophila melanogaster"

/mol_type="mRNA"

/db_xref="taxon:7227"

/clone="GH15518"

/sex="male and female"

/dev_stage="adult"

/lab_host="DH5 - alpha"

/clone_lib="GH Drosophila melanogaster head pot2"

/notes="Organ: head; Vector: pot2; Site 1: EcoRI; Site 2: XhoI; Sized fractionated cDNAs were directly ligated into pot2. Plasmid cDNA library."

ORIGIN

Query Match 22.0%; Score 331.2; DB 9; Length 607;
Best Local Similarity 76.7%; Pred. No. 3.1e-72;
Matches 405; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY 11 TGTTCGGCGCTTGGCGCTGCTGGCTTTGCTGCCGCTATCGGAGCAAGGTCTCTCACGAGA 70
DB 75 TGTTCCTCTGTTGATCTTCTGCGGATATTAAGAAAGCTGTCAAGGACCTCAAGAA 134
QY 71 AGAGACTCTTGAACGGCTTCTCGGAACTTACAAACCCCTGGAGCGGACCGGTGGCCAAAC 130
DB 135 AGCGCTCTGTGAACCATCTGCTCCACCTTACAATACGTCGAGCGGACCGGTGGCCAAATG 194
QY 131 AGAGCGACCGCTAGAGGTGAGTTCGGCTTGACCTTGCAGCAAAATCAATTCAGCTGACG 190

Db 195 AATCGGAGCCCTGGAGGTTAAAGTTTCGAGTGAAGCTGCAGCAGATCATCGACGTGGAGC 254
 QY 191 AGAAGATCAACTATTATAACCAATATATATGCTGTGTTGAGTGAATGACTACAACC 250
 Db 255 AGAAGATCAGCTTCTCATACGAATCTTTGGCTTTCTGTTGAGTGAACGACTACAATC 314
 QY 251 TGAGGTGGAACACACAGCAGTATGCGGGTCAAGGACCTCAGGATCACGCCCAACAAGT 310
 Db 315 TCGCTGGATGAACCGAATACGCGGGTCAAGGATCTAGCAATCACGCCCAACAAGC 374
 QY 311 TGTGAAGCCGACGCTCTTATGTATATAGTGTGCTGACGAGGTTTGTGACGGGACCTACC 370
 Db 375 TGTGAAGCCGACGCTCTCATGTACAACAGCGCGATGAGGATTCGATGSCAGTATC 434
 QY 371 AGACCAACGTGTGTGTCAGAACGCGGGGAGTTGCTGTACGTGACCACTGCAATATCA 430
 Db 435 ACACCAAGCTTGTGTCAACATACGCGGCGAGTTGTGTACGTGCGCCCTGTATCTTCA 494
 QY 431 AGAGCATATGCAAGATGACATCGGTGTTTCCCTTCGACGACCAACAACACTGTATGA 490
 Db 495 AGAGCATATGCAAGATGACATCACGTGTTCCCATTTGATGACCAACATTCGAAATGA 554
 QY 491 AGTTCGAGTGTGACATATGACGCAATCAGTTCGATCTGTTGCTAA 538
 Db 555 AATTCGAGTGTGACATGAGTGAATCAGTTCGATTTGTTGA 602

RESULT 3
 CD013901 2296 bp mRNA linear EST 21-OCT-2003
 LOCUS 90134548 single gene library Homo sapiens cDNA, mRNA sequence.
 DEFINITION
 ACCESSION CD013901
 VERSION CD013901.1 GI:37777431
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1. (bases 1 to 2296)
 Au-Yang, J., Fu, G.K., Wilson, A.D., Yang, J., Chien, D., Hawkins, P.R.,
 PCR isolation and cloning of novel splice variant mRNAs from known
 drug target genes
 Unpublished (2003)
 Contact: Jin, P.
 Incyte Corporation
 3160 Porter Drive, Palo Alto, CA 94304, USA
 Tel: 650 621 8639
 Fax: 650 621 8965
 Email: pjin@incyte.com.

FEATURES
 Location/Qualifiers
 1..2296
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="Single gene library"
 /note="Vector: pDrive Cloning Vector; RT-PCR was performed
 using gene-specific primers flanking the open-reading
 frame. PCR products were subcloned into pDrive Cloning
 Vector and sequenced completely using M13 forward and
 reverse primers. Sequencing gaps were closed by
 re-sequencing using primers flanking the gapped areas."

ORIGIN
 Query Match 20.1%; Score 302.2; DB 14; Length 2296;
 Best Local Similarity 56.9%; Pred. No. 1.2e-64;
 Matches 575; Conservative 0; Mismatches 433; Indels 3; Gaps 1;

QY 33 GCGTTTGTGCGCGATTCGAGCAAGGTCTCTCAGCAAGAGAGACTCTCTGAACGGGTGTCT 92
 Db 316 GGCATTGCGCGAGGAGGCTCGCATACCGAGACTGAGGACCGCTCTTCAAAACACCTCTT 375

QY 93 GCGCACTACACACCTCTGGAGCGCGTGGCCCAACGAGAGCAACCCCTAGAGGTGAC 152

Db 376 CCGGGCTACAAACCCCTGGCGCGCCGTCGCCCAACTTCAGACGTGTTGTCG 435
 QY 153 GTTCGGCTTGACCTTTGCAGAAATCAATTGACGTGCGAGAAAGAAATCAACTACTTATAAC 212
 Db 436 CTTTGGACTGTCCATCGCTCAGCTCATCGATGTGATGAGAAACCAATGATGACCCAC 495
 QY 213 CAATATATGCTGCTGTTGAGTGAATGACTACAACTGAGGTGGAACGACGAGTGA 272
 Db 496 CAACGCTCTGGCTTAAACACAGGAGTGGAGCGACTACAACTGCGCTGGAAACCCCGCTGATTT 555
 QY 273 TGGCGGGTCAAGGACCTCAGGATCACGCCCAACAAAGTTGTGAAACGCGAGCTCTTAT 332
 Db 556 TGGCAACATCATCTCTCAGGCTCCCTTCTGAGATGATCTGGATCCCGCAATTTCTT 615
 QY 333 GTATATATGCTGACGAGGGTTTTCAGCGGACCTTACAGACCAACGTTGGTGGTCAAG 392
 Db 616 CTACAACAATGCAGATGGGAGTTTTCAGTGAACCCACATGACCAAGGCCACCTCTTCTC 675
 QY 393 CGGCGGAGTTCCTGTAGTGCACCTGGGATATTCAGAGCACATGCAAGATGACAT 452
 Db 676 CACGGGCACTGTGCACTGGTGGTGGCGGCCATCTTACAAGAGCTCTCAGCATCGAGCT 735
 QY 453 CGCGTGGTTTCCCTTCGACGACCAACACTGTGATATGAGTTTCGTTAGTGCACATATGA 512
 Db 736 CACCTCTTCCCTTCGACGACGAGAACTGCAAGATGAAATTTGGTCTGAGCTTATGA 795
 QY 513 CGGCAATCAGTTGGATCTGGTCTAAAAGATGAGGCGCGCGGATCTATCGGACTTCAT 572
 Db 796 CAAAGCCAAAGATCGACCTGGAGCAGATGAGCAGACTGTGG--ACCTGAAGGACTACTG 852
 QY 573 AACAAATGGGAGTGGTATCTTAATAGGAATGCCAGGCAAAAGAAACACAAATACATACGC 632
 Db 853 GGAGAGCGCGAGTGGGCCATCGTCAATGCCACGGGCACTTACAACAGCAAGAGTACGA 912
 QY 633 GTGCTGCCCGGAGCCTTAGTGGAGCTCACCTTCCACCATCATGATGATAAGAAAGACCACTT 692
 Db 913 CTGCTGGCGGAGATCTACCCGAGCTCACCTACCGCTTTCGTCTATCGCGGCGCTGCCGT 972
 QY 693 GPACTACTTCTTCAACCTGATCGTCCGCTGGCTGTGATCTCATGATGGCACTCTCTCGG 752
 Db 973 CTTCTACACCACTCAACCTCATCATCCCTGCTGCTCATCTCTCTCTCTCTCTCTCTCT 1032
 QY 753 CTTTACACTGCCACAGACTCCGGAGAGAACTCACACTTGGAGTCACTATTTCTTCTATC 812
 Db 1033 CTTTCTACCTGCTCCGCTCCGAGTACGCTGTGATCACTGCTGTCTCTCTCTCTCTCT 1092
 QY 813 GCTGACGCTGTTCTCTCAACCTGGTAGCGAGACCTGCCACAGGCTCTCGAGCTATCC 872
 Db 1093 ACTCACCGTCTTCTCTGCTGCTCATCTGAGATCATCCGCTCCACTGCTGCTGCTATCC 1152
 QY 873 CTTGTAGGAGCGTACTTCAATTTGATCATGTTCTATGTTAGTGGTCTCTGTGGTACTGAC 932
 Db 1153 GCTCATCGCGAGTACCTGCTGTTTCAACATGATCTTGTGTCACCTGTCCATGCTATCAC 1212
 QY 933 TGTGTTGTTACTCAATTTACCACTCGAACAGCTGATATGATGAAATGCCACAGTGGAT 992
 Db 1213 CGTCTTGTGTTCAATTTGACACCGCTCCCGGAGCCACACCATGCCCACTGGCTGGT 1272
 QY 993 AAATCAGTATTTCTTCAATGTTGCCATGATGATCTCGGAATGTCGAGGCC 1043
 Db 1273 GCGGGGGGCCCTTCTGGGCTGTGTCGCCGCTTCTGATGAACCGGCC 1323

RESULT 4
 AL530299
 LOCUS AL530299 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
 DEFINITION cDNA clone CS0DD007YP05 5-PRIME, mRNA sequence.
 ACCESSION AL530299
 VERSION AL530299.2 GI:31068132
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

```
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 13, 2001 this sequence version replaced gi:12793792.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 91 91066 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
7646.r For more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0DD007CH03QP1&cluster=7646.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue Genoscope sequence ID : CS0DD007CH03QP1.
FEATURES
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DD007YP05"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 20.0%; Score 301.2; DB 9; Length 1201;
Best Local Similarity 60.2%; Pred. No. 1.6e-64;
Matches 537; Conservative 10; Mismatches 333; Indels 12; Gaps 3;
QY 65 ACAGAGAGAGACTCTGAGCGCTGTCGCGAACTACACACCCCTGGAGCGACCGGTGG 124
Db 197 WCCAGAGGAGCTTTTACAGGAGCTGGTCAAGAACTACAACTCCCTTGGAGAGCCCGTGG 256
QY 125 CCAACGAGACGCAACCGCTAGAGGTGAGCTTCGGCTTGACCTTGCAGCAATCATTTGACG 184
Db 257 CCAATGACTCGCAACCACTACCGTCTACTTCTCCCTGAGCTCTCGCATCATGACG 316
QY 185 TGGACGAGAGATCAACTACTTATTAACCAATATATGGCTGTCTGGTGGAGTGAATGACT 244
Db 317 TGGATGAGAAGAACCAAGTTTAAACCACCAACATTTGGCTGCAAAATGCTTTGACAGATC 376
QY 245 ACAACCTGAGGTGGAACGACGAGTATGGCGGGGTCAAGGACCTCAGGATCACGCCCA 304
Db 377 ACTATTACAGTGGAAATGTGCAAAATATCCAGGGGTGAAGACTGTTCGTTTCCCAAGTG 436
QY 305 ACAAGTTGTGGAAGCGGACGTCCTTATGTATATAGTGTCTGACGAGGGTTTTGACGGGA 364
Db 437 GCCAGATTTGGAACACAGACATTTCTTCTATAACAGTGTCTATGATGCGGCTTTGACGCCA 496
QY 365 CCTACAGACCAACGTTGGTGTGAGAGCGGCGGAGTGCCTGTGACGTGCCACCTGGCA 424
Db 497 CATTCACACTAAACGTTGTTGGTGAATTTCTTCTGGGCAATGGCCAGTWCCTGCCAGGCA 556
QY 425 TATTCAAGAGCATCAAGATGGACATCGCGTGGTTTCCCTTCGACGACCAACACTGTG 484
Db 557 TATTCAAGAGTTCCTGCTACATGATGACGTGTGTTCCCTTTGATGTGACACTGTGA 616
QY 485 ATATGAAGTTGGTAGTGGACATATGACGGCAATCAGTTGGATCTGGTCTAAAAAGATG 544
Db 617 AACTGAAGTTGGGTCTGCTCTACGGAGCTGGTCCCTTGGATCTGCAGATCCAGGA-- 674
QY 545 AGCAGCGCGCATCTATCGACTTCTATPACAAATGGGAGTGTGTATCTAATAGGAATGC 604
Db 675 -----GGCAGATATCAGTGGCTATATCCCAATGGAATGGGACCTAGTGGGAATCC 727
```

```
QY 605 CAGGCAAAAGACACAATAATACATACGCTGCTGCTCCCGAGCCCTAGCTGGAGCTCACCT 664
Db 728 CCGGCAAGAGGAGTGAAGGTTCTATGAGTCTKTCGAAAGAGAGCCCTACCCGATGTACCT 787
QY 665 TCACCATCATGATGATGAAGAAGACGAACCTTGTACTACTTCTTCAACCTGATCGTCCCGTGG 724
Db 788 TCACAGTGAACATGCGCCGACGAGACGCTCTACTATGGCCTCAACCTGCTGATCCCTGTG 847
QY 725 TGCTGATCTCATGATGGGACACTCCTCGGCTTACACTGCGCACAGACCTCCGAG--AGAA 782
Db 848 TGCTCATCTCCGCCCTCGCCCTGCTGTTCTCTGCTGCTGAGATTCGCGGGGAGAG 907
QY 783 ACTCACACTTGGAGTCACTATTCTTATCGCTGAGCGGTGTTCTCAACTGCTAGCCGA 842
Db 908 ATTTCCTCGGGATTAACAGTATTCTCTCTTTACCGTCTTATGCTGCTGCTGGCTGA 967
QY 843 GACCTTGCACACAGTCTCCGAGCGCTATCCCGCTTGTAGGAGCGTACTTCAATTGCAATCAT 902
Db 968 GATCATGCCCGCAACATCCGATTCGGTACCAATTGAWAGCCCACTTCTTCCGACAGCCAT 1027
QY 903 GTTCATGGTAGCGTCTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 954
Db 1028 GATCATCGT-GGGCTTTGGGTGGTTGKRMGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1078

RESULT 5
AY402873 1436 bp DNA linear GSS 15-DEC-2003
LOCUS Homo sapiens CHRNA3 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY402873
VERSION AY402873.1 GI:39758856
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1436)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1436)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment
FEATURES
source
1..1436
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1436
/gene="CHRNA3"
/locus_tag="HCWL1369"

ORIGIN
Query Match 19.6%; Score 294.8; DB 29; Length 1436;
Best Local Similarity 57.1%; Pred. No. 7e-63;
Matches 558; Conservative 0; Mismatches 417; Indels 3; Gaps 1;
QY 67 GAGAAGAGACTCTGAACCGTGTGTCGGGAACACTCAACACCCCTGGAGGACCGGTGGCC 126
```


Db 21 GAGCACCCTCTATTGTGAGCGGCTGTTTGAAGATTACAATGAGATCATCGGCGCTGTGCC 80
QY 127 AACGAGAGCGAACCCCTAGAGGTGAGGTTCGGCTTGACCTTCGACGAATCATTTGACGTG 186
Db 81 AACGTGTGTGACCCAGTCATCATCCATTTCCGAGGTGTCCATGCTCTCAGCTGGTGAAGGTG 140
QY 187 GACGAGAGAGTCACTACTTATTAACCAATATATATGCTGTCTGTTGAGTGGNATGACTAC 246
Db 141 GATGAAGTAACACGAGATCATGGAGACCAACCTGTGGCTCAAGCAATCTGGAATGACTAC 200
QY 247 AACCTGAGGTGGAACGACGAGTATGCGGGGTCAAGGACCTCAGGATCAGCGCCCAAC 306
Db 201 AAGCTGAAGTGAACCCCTCTGACTATGTTGGGCGAGATTATGCTGCTCCTCCACAG 260
QY 307 AAGTTGTGAAGCCGAGTCTCTTATGTAATATAGTGTGACGAGGTTTTCACGGGACC 366
Db 261 AAGATCTGAAGCCAGACATTTGCTGTATACAATGCTGTTGGGATTTCCAGGTGGAC 320
QY 367 TACCAGACCAACGCTGTGTGTCAGAGCGGCGCAGTTGCCGTGCTACGTGCCACTCGCAT 426
Db 321 GACAAGACCAAGCCTTACTCAAGTACACTGGGAGGTGACTTGGATACCTCCGCCCATC 380
QY 427 TTCAGAGCACATGCAAGATGAGATCGCGTGGTTTCCCTTCGACGACCAACACTGTGAT 486
Db 381 TTTAAGAGCTCCTGTAAATCGAGCTGACCTACTTCCCGTTTGATTACCAAACTGTACC 440
QY 487 ATGAATTCGGTGTGAGCATATGACGGCAATCAGTTGGATCTGGTGTAAAGATGAG 546
Db 441 ATGAATTCGGTGTGCTGCTCCTACGATAAGGCGAAATCGATCTGGTCTCTGA---TCGCG 497
QY 547 GCAGCGCGGATCTATCGGACTTCTAACAATGGGAGTGTATCTAATGAATGCCA 606
Db 498 TCTTCATGAACCTCAAGACTATTGGGAGCGCGAGTGGGCCATCATCAAGGCCCA 557
QY 607 GCGAAAAAGAACACATAACATACGCGTGTCTGCCCGAGCCCTACGTGACCTCACCTTC 666
Db 558 GGCTACAAACGACATCAAGTACAACTGCTCGGAGGAGATCTACCCGACATCAATAC 617
QY 667 ACCATCATGATGAAGAGAGAACCTTTGACTACTTCTCAACCTGANTGCTCCGTGCGTG 726
Db 618 TCGCTGTACATCGGCGGCTGCGCTTGTCTACACCATCAACCTCATCCTCCCTGCGCTG 677
QY 727 CTGATCTCATGATGCGACTCTCGGCTTCACTGCGCACAGACTCCGGAGAGAACTC 786
Db 738 ACCCTGTGCTATTTCTGCTCTCTCTCCCTGACGGTGTCTCTCTGCTGATCACTGAGACC 797
QY 847 CTGCCACAGGTCTCGACGCTATCCCGCTGTAGGAGCGTACTTCAATTTGCATCATGTTTC 906
Db 798 ATCCCTTCCACCTCGCTGTGCTATCCCGCTGTATGGAGAGTACCTCTGTTTACCACATGAT 857
QY 907 ATGGTAGCGTCTCTGTGTGCTACTGACTGTGGTGGTACTCAATTAACACCATCGAACAGCT 966
Db 858 TTGTGAACCTGTGCTCATGCTATCAACCGTCTCGTGTCTGCTGCTGCTGCTGCTGCTGCT 917
QY 967 GATATACATGAATGCCAGTGGATAAATCAGTATTCTCTCAATGTTGCTGCTGCTGCTGCTG 1026
Db 918 ACGACACACACAATGCCCTCATGGGTGAAGACTGTATTCTTGAACCTGCTCCCGAGGCTC 977
QY 1027 CTGCGAATGTCAGGGCCA 1044
Db 978 ATGTTTCATGACCAAGGCCA 995

RESULT 6

AK034228

LOCUS

DEFINITION

2940 bp mRNA linear HTC 18-SEP-2003
Mus musculus adult male diencephalon cDNA, RIKEN full-length
enriched library, clone:933016516 product:cholinergic receptor,
nicotinic, alpha polypeptide 4, full insert sequence.

ACCESSION

AK034228.1 GI:26329798

VERSION

HTC; CAP trapper.

KEYWORDS

Mus musculus (house mouse)

SOURCE

Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

1 Carninci, P. and Hayashizaki, Y.

AUTHORS

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

TITLE

High-efficiency full-length cDNA cloning
prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL

Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE

20499374

PUBMED

11042159

REFERENCE

3

AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsunai, T., Tashiro, H., Itoh, M.,
Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujisake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, S., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

TITLE

RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multipillar sequencer

JOURNAL

Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE

20530913

PUBMED

11076861

REFERENCE

4

AUTHORS

The RIKEN Genome Exploration Research Group Phase II Team and the

TITLE

Functional annotation of a full-length mouse cDNA collection

JOURNAL

Nature 409, 685-690 (2001)

REFERENCE

5

AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research

TITLE

Group Phase I & II Team.

JOURNAL

Analysis of the mouse transcriptome based on functional annotation

REFERENCE

of 60,770 full-length cDNAs

JOURNAL

Nature 420, 563-573 (2002)

REFERENCE

6

AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akashira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A.,
Muramatsu, M. and Hayashizaki, Y.

TITLE

Direct SubMISSION

JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216]

COMMENT

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/.

FEATURES
Location/Qualifiers
Source

```

1. .2940
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:9330165116"
/db_xref="MGI:2398424"
/db_xref="taxon:10090"
/clonetype="9330165116"
/sex="male"
/tissue_type="diencephalon"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_status="adult"
116. 2005
/notes="unnamed protein product; cholinergic receptor,
nicotinic, alpha polypeptide 4 (MGD|MGI:87888,
GB|NW_015730, evidence: BLASTN, 99%, match=1946)
putative"
/codon_start=1
/protein_id="BAC28638.1"
/db_xref="GI:26329799"
/translation="MEIGSGAPPLPPLPLPLLLLTGTLPPASSHLETHAEERLLK
RLEPGNKNRSPVANISDVLRPLSLIAQLIDVDENQMTINVMKQNDHDKLRW
DPEDYENVTISIRIPSELIWRPDLVLYNADGDFAVTHLKAHLFYDGRVQWTPPALYK
SSCIDVTFFPDQDQTKWGSWTYDKAKIDLVSMSRVQDLDFWESGEWIVDAVG
TNTKRYECCABTYPIITVAFIIRRLPLFTINLIIPCLLISCLTFLVFLPSECGEK
VTLCSVLSLAVFLLLIETIIPSTSLVPLIGELVLLFTMIFTVLSIVITVFLNVHH
RSPRHTMPAWRVRFLDIVRPLLFKRVSVKDCRRLLKESMHWANAPRPWPPEP
EGLIGLIDCNQSLSPAPTCNMRDIAVETQTCRSPSHKVPDLKTSVEKSPCPSPG
SCHPNISGAPVLLIKARLSQVSPSSBAEGRSICRSRSIQYCVSDGAASLSTESK
PTGSPASLTRESQVSPDQSPCKTCKEPPSPITVLKAGTKAPQHLPLSPAL
TRAVEGVQYIADHLKAEIDFVSKEDWKVAVVIDRIFLWMPFIIVCLLTGVLGLPPW
LAGMI"

```

CDS

Query Match 19.5%; Score 293.2; DB 11; Length 2940;

Best Local Similarity 55.9%; Pred. No. 2.5e-62;

Matches 578; Conservative 0; Mismatches 453; Indels 3; Gaps 1;

```

QY 11 TGTGGCGGCTTGGCGCTTGGCTTGTGTCGCCGATCGGAGCAAGGTCTCTCAGAGA 70
Db 174 TCTTAGGACCGGCTTGTGCTGTAGGACGACATAGAGACCCGGGCGCATCGCGAGG 233
QY 71 AGAGACTCTCTGAACGGTGTGTCGGCAATACAAACCTCGGAGCGACCGGTGGCCAAAG 130
Db 234 AGCGGCTCTGAAGAGACTTCTCTGGCTACACAAAGTGTCTCGCCAGTAGCCAAATA 293
QY 131 AGAGCAACCGCTAGAGGTAGTTCGGCTTGACCTTGACGAGCAAAATCATGTAGCGGAG 190
Db 294 TCTCAGATGTGTCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 353
QY 191 AGAAGAATCAACTACTTATAACCAATATATGCTGTGCTGAGTGGATGACTACCAACC 250
Db 354 AGAANAACAGATGATGACGACCAACGCTGTGGTGAAGCAGAGTGGCATGACTACAAAC 413
QY 251 TGAGGTGGACGACGAGTATGCGGGGTCAAGGACCTCAGGATCAGCCCAACCAAGT 310
Db 414 TCGCGTGGACCTGTGCTGACTACGAGAAATGTCACTCCATCGCATCCCATCTGAATCA 473
QY 311 TGTGNAAGCGGACGCTCTTATATATAGTGTGCTGACGAGGGTTTGTAGGGACCTACC 370
Db 474 TCTGGAGGCTGACATCGCTCTTACAAACGCGGACGGGACCTTGTGAGTCAACCCACC 533
QY 371 AGACCAACGTTGGTGTGAGAGCGGCGGAGTGTGCTGTACGTGCCACCTGGCATATCA 430
Db 534 TAACCAAGCCACCTGTCTATGATGGCGGTGTGAGTGGACACCCCGGCACTATA 593
QY 431 AGAGCAATGCAAGATGAGATCGCTGTTTCCCTTCGACGACCAACCACTGTGATGCA 490
Db 594 AGAGCTCTGCAATGAGATCGCTGTTTCCCTTCGACGACCAACCACTGTGATGCA 653
QY 491 AGTTCGGTAGTGGACATATGAGCGCAATCAGTTGGATCTGTGCTGCTGCTGCTGCTGCT 550

```

ORIGIN

```

Db 654 AGTTTGGGTCTTGAGCCTACGACCAAGGCTGACCTTGGTGGATGCACAGCGCGT 713
QY 551 GGGCGCATCTATCGGACTTTCATACAAATGGGAGTGGTATCTAAATAGGAATGCCAGGCA 610
Db 714 TGGACCAACT--GGACCTTCTGGGAAAGTGGGAGTGGGTCAATTTGGATGCGCGTGGGCA 770
QY 611 AAAAGAAACAATAAATACATAGCGTGTGCCCCGAGCCCTACGTGGGAGCTCACCTTCCACCA 670
Db 771 CCTACCAACACAGGAGTATGAATGTGTGCCGAGATCTATCTGATCATCACCCTACGCCT 830
QY 671 TCATGATAGAGAGAGAACTTGTACTACTTCTTCAACCTGATGTCCTCGTGGTGGTCTGTA 730
Db 831 TCATCATCCGCGACTGCGCACTGTTCTACACCATCAACCTTATCATCCCGTGGCTGCTCA 890
QY 731 TCTCATGATGGCACTCTCGGCTTCCACACTGCTCCAGAGACTCCGAGAGAACTCACAC 790
Db 891 TCTCTGCTCTACCGGTGGTCTTCTATCTGCGGTGGGAGTGGCGGAGAGGTCAACGC 950
QY 791 TTGGAGTCACTATTTCTTATCGCTGACGCTGTCTCAACCTGTGAGCGGAGACCTGTC 850
Db 951 TGTGATCTCGGTGCTGCTTCTCTACCGTCTTCTGCTGCTCATCCGAGATCATCC 1010
QY 851 CACAGGTCTCCGAGCGTATCCCGCTGTAGGAGCGTACTTCAATTCGATCATGTTCTATGG 910
Db 1011 CGTCCACCTCGGTGCTATCCCGCTCATCGGCGAGTACCTGCTCTTCCACCATGATCTTCG 1070
QY 911 TAGCGTGTCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 970
Db 1071 TCACCTCTCTCAATTTGTCATCGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1130
QY 971 TCATGAAATGCCACAGTGGATAAAATCAGTATTCCTACAAATGTTGGCTGCTGCTGCTGCT 1030
Db 1131 CACACACCATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1190
QY 1031 GAATGTCGAGGCA 1044
Db 1191 TCATGAAGCGGCA 1204

```

```

RESULT 7
AK083157
LOCUS
DEFINITION
Mus musculus adult male hippocampus cDNA, RIKEN full-length
enriched library, clone:C630019M18 product:cholinergic receptor,
nicotinic, alpha polypeptide 4, full insert sequence.
ACCESSION
AK083157
VERSION
AK083157.1 GI:26350296
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE
99279253
PUBMED
10349636
REFERENCE
2
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
JOURNAL
20499374
MEDLINE
11042159
PUBMED
11042159
REFERENCE
3
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsumi,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwara,K., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanabe,M.,

```

Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000) 20530913 PUBLISHED 11076861

REFERENCES

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)

AUTHORS

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002) 6 (bases 1 to 3230)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akashira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

Direct Submission

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.go.jp/>

URL: <http://fantom.gsc.riken.go.jp/>

Location/Qualifiers

1. 3230

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM_DB:C630019M18"

/db_xref="MGI:2417536"

/db_xref="taxon:10090"

/clone="C630019M18"

/sex="male"

/tissue_type="hippocampus"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

36. 1925

/note="unnamed protein product; cholinergic receptor,

nicotinic, alpha polypeptide 4 (MGI:MGI:87888,

GB|NM_015730, evidence: BLASTN, 99%, match=1946)

putative"

/codon_start=1

/protein_id="BAC38788.1"

/db_xref="GI:26350297"

/translation="MEIGSGAPPPLLLIPLLILGTLGPASSHETRAHEERLLK

RLPSGNKSRVPISELVIRPGLSLAQILVDKQNMFTNWWVQOEWDYKLRW

DPGDYNTSRIPSELVIRPGLSLAQILVDKQNMFTNWWVQOEWDYKLRW

SSCSIDVTFPPDQNCMTKFGSNTYDKAKIDLVSMSRVDQLDFWESGEWIVDAVG

TYNTRKYCCCAIYDPIVAVTIRRLPLFYTLNLIIPCLISLTLYLPLSPGCEK

ORIGIN

Query Match 19.5%; Score 293.2; DB 11; Length 3230;
Best Local Similarity 55.9%; Pred. No. 2.7e-62;
Matches 578; Conservative 0; Mismatches 453; Indels 3; Gaps 1;
LAGMI"

QY 11 TGTGGCGCGCTTGGCGCTGCTGGCTTGTGTCGGCTTGCAGCAAAATCATTTGACGTGACG 70
Db 94 TCTTAGGACCGCGCTTGTGCTGCTAGCAGCCACATAGAGACCGCGGCCCATGCGGAGG 153
QY 71 AGAGACTCCTGGAACGGCTTGTGCTGGCGAACTACACACCCCTGGAGCGACCGGTGGCAACG 130
Db 154 AGCGGCTCCTGAAGAGACTTCTCTGCGGTACACCAAGTGGTCTCGGCCAGTAGCAATA 213
QY 131 AGAGCGAACCGCTTAGAGGTGAGGTTCGGCTTGCACCTTGCAGCAAAATCATTTGACGTGACG 190
Db 214 TCTCAGATGGTCTTGTGCTGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 273
QY 191 AGAGAAATCAACTACTTATTAACCAATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 250
Db 274 AGAAAAACAGATGATGACGACCAACGCTGGGTGAAGCAGGAGTGGCATGACTACAAAC 333
QY 251 TGAGGTGGAAACACAGCAGGAGTATGGCGGGTCAAGGACCTCAGGATCAGCCCAACAAGT 310
Db 334 TCGCTGGGACCGCTGGTGAATGACAGAAATGTCACTCCATCCGATCCCATCTGAACTCA 393
QY 311 TGTGAAGCGCGACGCTCTTATGATATATAGTGTGCTGACGAGGCTTTTACGCGGACCTTACC 370
Db 394 TCTGGAGGCTGACATGCTCTTACACACACGCGACGCGGAGCTTTGAGTCACTCACCACC 453
QY 371 AGACCAACGTGTGTGTGACAGCGCGGCGAGTTCCTGTGACCTGCGACCTGGCATATTTCA 430
Db 454 TAAACAAACCCACCTGTTCTATGATGGCGGTGTGCACTGGACACACCCCGGCATCTATA 513
QY 431 AGAGCACATGCAAGATGGACATCGGTGTTTCCCTTCGACACCAACACTGTGATATGA 490
Db 514 AGAGCTCTTCGACGATCGACGTCCTTCTTCCCTTCGACGAGAACTGTACCATGA 573
QY 491 AGTTCGCTAGCTGGACATATGACGCAATCAGTTGATCTGCTGTCTGCTGCTGCTGCTGCTG 550
Db 574 AGTTTGGGTCTGGACCTACGACAGGCCAAAGTTGACTTGGTGGATGTCACAGCCGTG 633
QY 551 GCGGCGATCTATCGGACTTCTATAAATAATGGGAGTGGTATCTAATAGGAATGCCAGGCA 610
Db 634 TGGACCAACT---GGACTTCTGGGAAAGTGGGAGTGGGTCTTGTGGATGCGGTGGCA 690
QY 611 AAAAGAACACATAACATACGCTGCTGCGCCGAGCCCTACGTGACGCTCACTTCACTCA 670
Db 691 CCTACACACAGGAAGTATGAATCTGTGCGGAGATCTATCTGACATCACTGACCTAGCCT 750
QY 671 TCATGATAAGAAAGACCAACTTGTACTTCTTCAACTGATGCTGCTGCTGCTGCTGCTGCTG 730
Db 751 TCATCATCGCGACTGCCACTGTTCTACACCACTATCATCCGTCGCTGCTCA 810
QY 731 TCTCATGATGCACTCTCTCGGCTTCACTGCCACCAAGACTCCGAGAGAACTCACAC 790
Db 811 TCTCTGCTCACCCTGCTGCTTCTTCTCTTCTCACCCTGCTTCTGCTGCTGCTGCTGCTG 870
QY 791 TTGGAGTCACTATCTTCTATGCTGACGCTGCTTCTCACTGATGCTGCTGCTGCTGCTGCTG 850
Db 871 TGTGATCTCGGTGCTGCTTCTTCTCACCCTGCTTCTGCTGCTGCTGCTGCTGCTGCTG 930
QY 851 CACAGGTCTCGGACCTATCCCTGTTAGGAGCACTTCTCAATTTGATCATGTTTCATGG 910
Db 931 CGTCCACCTCGTGTGCTATCCGCTCATCGGAGTACTGCTCTTCACTGATGTTCTCG 990

FEATURES

source

CDS

```
QY 911 TAGCGTCGCTCTGTGGTACTGACTGTGGTGTACTCAATTACCACCATCGAAAGCTGATA 970
Db 991 TCACCTCTCCATTGTCATCAGGTCTTCGTCTCAATGTACACACCGCTCACCAGCA 1050
QY 971 TACATGAATGACACAGTGGATATAATCAGTATTCCTCAATGGTGGCATGATCTGC 1030
Db 1051 CACACACCATGCGCGCTGGGTGCGCAGAGTCTTCCTGGACATGTGCGCCGCTCTCTCT 1110
QY 1031 GRATGTCGAGGCCA 1044
Db 1111 TCATGAAGCGCCA 1124

RESULT 8
AY402874 1436 bp DNA linear GSS 15-DEC-2003
LOCUS Pan troglodytes CHRNA3 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY402874
VERSION AY402874.1 GI:39758857
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 1436)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrieri,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1436)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrieri,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source 1..1436
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>1436
/gene="CHRNA3"
/locus_tag="HCMI369"

ORIGIN
Query Match 19.1%; Score 286.8; DB 29; Length 1436;
Best Local Similarity 55.6%; Pred. No. 7.3e-61;
Matches 544; Conservative 0; Mismatches 431; Indels 3; Gaps 1;

QY 67 GAGAGAGACTCCTGAACGGTCTCTCGGCACTACACACCCCTGGAGCGCGTGGCC 126
Db 21 GAGCAGCGTCTATTGTAGCGGCTGTTTGAAGATTACAAATGAGATCATCCGGCCTGTGGCC 80
QY 127 AACGAGAGCGAACCGCTAGAGGTGAGGTTCGGCTTGACCTTGCAGCAAAATCAATGACGTG 186
Db 81 AACGTGCTGACCCAGTCATCATCAATTTNNNGTGTCCATGTCTNNNTGGTGAAGTG 140
QY 187 GACGAGAAGATCACTACTTATTAACCAATATATGGCTGTGCTGGAGTGAATGACTAC 246
Db 141 GATGAAGTAAACCAAGATCATGGAGACCAACNNGNCTCAAGCAAAATCTGGAATGACTAC 200
QY 247 AACCTGAGGTGGAACGACGAGGTATGCGGGGTCAAGGACCTCAGGATCACCCCAAC 306
Db 201 AAGCTGAAGTGAACCCCTCTGACTATGGTGGGCGAGGTTCATGCTGCTCCCTGNNNN 260
```

```
QY 307 AAGTTTGGAAAGCCGACGCTCCTTATGTATAATAGTGTGCTGACGAGGTTTTCAGCGAACC 366
Db 261 NNNNNCTGGAAGCCNNACATTTGCTGTATTAACAATGCTGTTGGGATTTCCAGGTGGAC 320
QY 367 TACAGACCAAGTGTGTGTGTCAGAGCGGCGGAGTTCCTGTACGTGTCACCTGSCATA 426
Db 321 GACAAGACCAAAAGCCTTACTCAAGTACACTGGGGAGGTGACTTGGATACCTCCGGGCATC 380
QY 427 TTCAAGAGCACATGCAAGATGGACATCGCGTGTGTTTCCCTTCGACGACCAACAACCTGTGAT 486
Db 381 TTTAAGAGCTCCTGCAAAATCGACGTGACTCTTCCCGTTTGATTACCAAAACTGTACC 440
QY 487 ATGAAGTTCGTAGCTGGACATATGACGGCAATCAGTTGGATCTGTGTCTGTGCTTAAAGATGAG 546
Db 441 ATGAAGTTCGTGTTCTGTTCTTACGATAAAGGGCAAAATCGATCTGTGCTCTGA---TCGGC 497
QY 547 GAGGGGGGATCTATCGGACTTCATAACAATGGGAGTGGTATCTAATAGAAATGCCA 606
Db 498 TCTTCCATGAACCTCAAGGACTATTGGGAGAGCGGAGTGGGCCATCATCAAGGCCCA 557
QY 607 GGCAAAAGAACACAATAACATACGCGTGTGCTGCCCGAGCCCTACGTGGACGTCACCTTC 666
Db 558 GGTCTACAAACAGACATCAAGTACAACTGCTNNNAGAGATCTACCCCGACATCACATAC 617
QY 667 ACCATCATGATGAAGAAGACGAACTTTGTACTACTTCTTCAACCTGATGCTCCCGTGGTG 726
Db 618 TCGCTGTACATCGCGGCGCTGCTGTTTCTACACCATCAACCTCATCATCCCTGCTGCTG 677
QY 727 CTGATCTCATGATGACACTCCTCGGCTTCACACTGCCACGACTCCGGAGAGAACTC 786
Db 678 CTCATCTCTCTCTACTGTCGTCTTCTACTCTGCCCTCCGACTCGGTGAGAGGTG 737
QY 787 ACACTTGGAGTCACTATTCTTCTATCGTACGCGTGTCTCTCAACCTGGTAGCGGAGACC 846
Db 738 ACCCTGTGCAATTTCTGCTCTCTCTCTCTGACGGTGTCTCTCTGGTGTGATCAGTACAGCC 797
QY 847 CTGCCACAGGTCTCCGACGCTATCCCGCTGTTTAGGAGACGTACTTCAATTCGATCATGTTTC 906
Db 798 ATCCCTTCCACCTCGCTGGTGTGATCCCGCTGATTGGAGAGTACCTCTCTGTTTCCACATGATT 857
QY 907 ATGGTAGCTGCTGCTGTGGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 966
Db 858 TTTGTAAACCTGTCCATCGCTATCGCTATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 917
QY 967 GATATACATGAAATGCCACAGTGGATAAAATCAGTATTCCTACAATGTTGCTCATGGATA 1026
Db 918 ACACACACACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 977
QY 1027 CTCGGAATGTCGAGGCCA 1044
Db 978 ATGTTTCATGACGAGGCCA 995
```

```
RESULT 9
AK081254 3483 bp mRNA linear HTC 20-SEP-2003
LOCUS Mus musculus adult male corpus striatum cDNA, RIKEN full-length
DEFINITION enriched library, clone:030030P04 product:cholinergic receptor,
nicotinic, beta polypeptide 2 (neuronal), full insert sequence.
ACCESSION AK081254
VERSION AK081254.1 GI:26099790
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1
AUTHORS Carninci,P. and Hayashizaki,Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
```

```

REFERENCE
AUTHORS
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL
MEDLINE
20499374
PUBMED
11042159
REFERENCE
AUTHORS
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuami,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL
MEDLINE
20530913
PUBMED
11076861
REFERENCE
AUTHORS
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE
Functional annotation of a full-length mouse cDNA collection
JOURNAL
MEDLINE
Nature 409, 685-690 (2001)
REFERENCE
AUTHORS
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL
MEDLINE
Nature 420, 563-573 (2002)
REFERENCE
AUTHORS
6 (bases 1 to 3483)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Saitoh,K., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tanaka,A., Takahashi,P., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
TITLE
Direct Submission
JOURNAL
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
COMMENT
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.
FEATURES
source
Location/Qualifiers
1..3483
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM,DB:C030030P04"
/db_xref="MGI:2413538"
/db_xref="taxon:10090"
/clone="C030030P04"
/sex="male"
/tissue_type="corpus striatum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"

```

```

/dev_stage="adult"
1..3483
/notes="cholinergic receptor, nicotinic, beta polypeptide 2
(neuronal) (MGDI:87891, GB|NM_009602, evidence: BLASTN,
99%, match=1498)"
ORIGIN
Query Match 18.7%; Score 281.6; DB 11; Length 3483;
Best Local Similarity 57.7%; Pred. No. 2.3e-59;
Matches 547; Conservative 0; Mismatches 389; Indels 12; Gaps 2;
QY 99 CTACAAACCCCTGGAGCGCGGTGGCCCAACAGAGAGCGAAGCGCTAGAGCTCAGGTTCGG 158
Db 326 CTATATAAAGCTGATCCGTCAGCTACTAATGGCTCTGAGCTGGTACTGTACAGCTCAT 385
QY 159 CTTGACCTTTCAGCAATCAATCACTGACGTGACGAGAGAAATCACTACTTATTAACCAATAT 218
Db 386 GGTATCATTTGGCACAGCTCATCAGTGTGCACGCGGAGCAGATCATGACCAACGCT 445
QY 219 ATGGCTGTCTGGAGTGGAACTACAACTGAGGTGGAACGACGAGCGAGTATGGCGG 278
Db 446 CTGGCTGACCCAGGAGTGGGAAGATTATCGCCTCACATGGAAGCCTGAGGATTCGACAA 505
QY 279 GGTCAAGGACCTCAGGATCAGCCCAACAAAGTTGTGGAAGCGGAGCTCTTATGTATAA 338
Db 506 TATGAAGAAGTCCGACTCCCTTCTTAAGCACATCTGGCTCCAGATGTGGTCTTATACAA 565
QY 339 TAGTGTCTGACGAGGGTTTTGACGGGACCTACAGACCAACGCTGGTGGTTCAGAAAGCGCGG 398
Db 566 CAATGTGACGGCATGTACGAAGTCTCTTCTTATTCGAATGCTGTGTCTCTCTATGATGG 625
QY 399 CAGTTCCTGTAGTGGCCACTGGGCATATTCAAGAGCACATGCAAGATGAGACATCGCGTG 458
Db 626 CAGCATCTTTTGGCTACCGCTGCCATCTACAGAGCGCATGCAAGATGAGGTGAAGCA 685
QY 459 GTTTCCTTCGAGGACCAACACTGTGATGATGAAGTTGGTGGTGGTGGTGGTGGTGGTGG 518
Db 686 CTTCCCATTTGACGACGACACTGACACCATGAAGTTCCGCTCTCGGACCTACGACCGCAC 745
QY 519 TCAGTTGGATCTGGTCTAAAGATGAGGACGAGCGCGCGATCTATCGGACTTCTATAACAA 578
Db 746 TGAGATTGACCTGGTCTCAAAAGCGATGTGGCGAC--CTGGACGACTTTCACACCTAG 802
QY 579 TGGGAGTGGTATCTAATAGGATGCGAGGACAAAGAACACATACATACCGCTGCTG 638
Db 803 TGGGAGTGGGACATCATCGCACTGCCAGCGCGACGACCAACGAGACCCAGACGACTCC-- 860
QY 639 CCGGAGCCCTACGTGGAGCTCACCCTTACCATCATGATAAGAGACGAACTTGTACTA 698
Db 861 -----ACTAGTGGACATCACCTACGACTTCATCTCGCAAAACCGCTCTCTA 913
QY 699 CTTCTTCAACCTGATCGTCCGCTGGCTGTGATCTCATCGATGGCACTCTCGGCTTCAC 758
Db 914 CACCATCAACCTCATCATCCCTGGCTACTCATCACTCGCTGGCCATCCTGCTCTTCTA 973
QY 759 ACTGCCACAGACTCCGGAGAGAACTCACACTTGGAGTCACTATTCTTCTATCGCTGAC 818
Db 974 CTTGCCCTCAGACTGTGGTGAAGAATGACACTTTGTATTCTGTCTGTGGCGCTAC 1033
QY 819 GGTGTTCTTCAACCTGGTAGCGAGACCTGCCACAGGCTCTCGACGCTATCCCCCTGTT 878
Db 1034 GGTGTTCTTCTGTGCTCATCTCAAGATTGTGCTCCACCTCTCCAGAGTACCGTGGT 1093
QY 879 AGGAGCGTACTTCAATTGTCATGTTTCATGGTAGCGTCTGTGGTACTGACTGTGGT 938
Db 1094 GGGAAAGTACCTCATGTTTCAACCATGGTGTAGTTCACCTTCTCATGCTCCTAGCGGTG 1153
QY 939 GGTACTCAATTACCAACCATCGAACAGCTGATATATGAATGAATGCCACAGTGGATAAATC 998
Db 1154 TGTGCTCAATTGTCACACCGCTTGGCTTACCACGACACCATGGGCCCTCGGTCAAGGT 1213
QY 999 AGTATTCTTACATGGTTCCTGCGATGATGCGGAATGTGAGGCCAGG 1046

```


QY 519 TCAGTTGGATCTGGTCTAAGAGATGAGCGGCGGATCTATCGGACTTCAACAAA 578
Db 742 TGAGATTGACCTGGTGCTCAAGAGCATGTGGCCAGC---CTGGAGCATTTACACCTAG 798
QY 579 TGGGGAGTGGTATCTAATAGGAATGCCAGGCAAAAAGAACACAAATACATACGCGTGCTG 638
Db 799 TGGGGAGTGGGACATCATCGCATGCCAGCGGCGACGAGACCCAGAGCACTCC-- 856
QY 639 CCCGAGCCCTAGGTGAGTGCACCTTACCATCATGATGAAGAGACGACCTTGTACTA 698
Db 857 -----ACCTAGCTGGACATCACCTACGACTTCATCATTCGTGCGAAACCGCTCTTCTA 909
QY 699 CTTCTTCAACCTGATCGTCCCGTGGCTGTATCTCATGATGCGCACTCTCGGCTTAC 758
Db 910 CACCATCACTCATCATCCCTGGGTACTCATCATCTCGTGGCCATCTGTCTTCTTA 969
QY 759 ACTGCCACAGACTCCGGAGAGAACTCACACTTGGAGTCACTATTCTTCTATCGCTGAC 818
Db 970 CTGCGCCTCAGACTGTGGTGAAGATGACACTTTGTTTCTGTCTGTGGCGCTAC 1029
QY 819 GGTGTCTTCACTCGTGGTGGAGACCTGCCACAGGCTCTGGAGCGTATCCCGCTGTT 878
Db 1030 GGTGTCTTCTGCTCATCTCCAAAGATTGGCTCCCACTCTCCCTCGAGTACCGTGT 1089
QY 879 AGGGAGCTACTTCAATTGTCATCATGTTCAATGGTGGTGTCTGTCTGCTACTGACTGTGT 938
Db 1090 GGAAGTACCTCATGTTTACCATGTTGTAGTACCTTCTTCATCGTCACTAGCGTGTG 1149
QY 939 GGTACTCAATTACCACTCGAACAGCTGATATACATGAATGCCACAGTGGATAAATC 998
Db 1150 TGTGCTCAATGTGCACACCGTTGCGCTACCAGCACCATGGCGCTGGTCAAGT 1209
QY 999 AGTATCTTCAATGTTGCCATGGATACCTGGAATGTGAGGCCAGG 1046
Db 1210 GGTCTTCTGGAGAGCTGCCACCTCTCTTCTTCTGACGACGCCACG 1257

RESULT 11
AK051742
LOCUS
DEFINITION
Mus musculus 12 days embryo spinal ganglion cDNA, RIKEN full-length
enriched library, clone: D130070121 product: cholinergic receptor,
nicotinic, beta polypeptide 2 (neuronal), full insert sequence.
AK051742
AK051742.1 GI:26342173
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049

ORIGIN	AK053497	1864 bp	mrna	linear	HTC 20-SEP-2003
Query Match	18.7%;	Score 281.6;	DB 11;	Length 4046;	
Best Local Similarity	57.7%;	Pred. No. 2.5e-59;			
Matches 547; Conservative 0;	Mismatches 389;	Indels 12;	Gaps 2;		
QY	99	CTACAAACCTCGAGCGACCGGTGGCCAAACGAGAGCGAAACCGCTAGAGTTCAGGTTCCG	158		
DB	331	CTATAACAAGCTGATCCGTCAGCTACTAATGGCTCTGAGCTGGTGAAGTCTACAGGTCAT	390		
QY	159	CTTGACCTTCGACGAATCATTTACGTGGGACGAGAGAAATCAACTACTTATACCAATAT	218		
DB	391	GGTATCATTGGCACAGCTCATGCTGTGCACGCGGAGCAGATCATGACCAACCAAGT	450		
QY	219	ATGGCTGTCTGAGTGGGAATGACTACAACCTGAGGTGGAACGACAGCGAGTATGGCGG	278		
DB	451	CTGGCTGACCCAGAGTGGGAAGATTATCGCTTCACATGGAGCCTGAGGATTCGACAA	510		
QY	279	GGTCAAGGACCTCAGGATCAAGCCCAACAAAGTTGTGGAACCGGAGCTCTTATGATATAA	338		
DB	511	TATGAAGAAAGTCCGACTCCCTTTAAGCACATCTGGCTCCAGATGTGTTCTATACAA	570		
QY	339	TAGTGTGACGAGGGTTTTCACGGACCTACAGACCAAGCTGGTGGTACAGAGCGCGG	398		
DB	571	CAATGCTGACGGCATGTACGAAGTCTCTTCTATTCCAATGCTGTGGTCTCCCTATGATGG	630		
QY	399	CAGTTGCTGTACGTGCCACTGGCATATTTCAAGAGCACATGCAAGATGACATCGCGTG	458		
DB	631	CAGCATCTTTTGGCTACCGCTGCCATCTACAGAGCGCATGCAAGATTGAGGTGAAGCA	690		
QY	459	GTTCCTCTCGACGACCAACACTGTGTGATGAAGTTGGTGGTGGACATATGACGGCAA	518		
DB	691	CTTCCCATTTGACGACGAGAACTGCACATGAAGTTCCGCTCTCGGACCTACGACCGCAC	750		
QY	519	TCAGTTGGATCTGGTGTAAAGATGAGGACGCGGCGATCTATCGACTTCTATACAAA	578		
DB	751	TGAGATTGACTGGTGTCTAAAGCGGATGTGGCCAGC---CTGAGCACTTCACACCTAG	807		
QY	579	TGGGGAGTGTATCTAATAGGAATGCCAGGCAAAAGAACACATACATACATCGGTGCTG	638		
DB	808	TGGGGAGTGGACATCATCGCATGCGCAGCGCCGACGACGAGAACCCAGCACTCC--	865		
QY	639	CCCGAGCCCTAGTGGACGTCACTTCCATCATCATGATGAAGAGCAACCTTGTACTA	698		
DB	866	-----ACCTAGGTGGACATCACTACGACTTTCATCATTCGTGCGAAACCGCTCTCTA	918		
QY	699	CTTCTTCAACTGATCGTCCGGTGGCTGATCTCATCGATGCACTCTCGCTCGCTTCAC	758		
DB	919	CACATCAACTCATATCCCTCGGTACTCATCACTCTGCGCCATCCTGCTGCTTCTA	978		
QY	759	ACTGCCACCACTCCGAGAGAACTCACACTTGGAGTCACTATTCTTCTATCGCTGAC	818		
DB	979	CCTGCCCTCAGCTGGTGGAAAGATGACACTTTGTATTCTCTGCTGCTGGCGCTCAC	1038		
QY	819	GGTGTCTCTCAACTGGTAGCGAGACCTTCGACAGGTTCTCGACGCTATCCCCCTGTT	878		
DB	1039	GGTGTCTCTCTGCTCATCTCCAAAGATTGTGCTCCCACTCCCTCGACGTAGCGGTGT	1098		
QY	879	AGGAGCTACTTCAATTGCATCATGTTTCATGGTAGCTGCTCTGTGTTACTGACTGTGGT	938		
DB	1099	GGGAAGTACTCATGTTTACCATTGGTGTAGTCACTCTTCCATCTGCTACTAGCGGTG	1158		
QY	939	GGTACTCAATTACCACCATCGAAGCTGATATACATGAATGCCAGTGGGATAAAATC	998		
DB	1159	TGTGCTCAATGTGCACACCGTTTCCGCTACCCAGCACACCATCGGCGCTGGGTCAAGGT	1218		
QY	999	AGTATTCTCAATGGTTGGCATGATGATCTGCGAATGTTCGAGGCGAGG	1046		
DB	1219	GGTCTTCTCGGAAAGTGTGCCCACTCTCTCTCTCTGTCAGCAGCCAG	1266		
LOCUS DEFINITION	Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched library, clone:EL130103E14 product:NEURONAL NICOTINIC ACETYLCHOLINE RECEPTOR ALPHA 3 SUBUNIT, full insert sequence.				
ACCESSION	AK053497.1 GI:26343494				
VERSION	HTC; CAP trapper.				
KEYWORDS	Mus musculus (house mouse)				
SOURCE	Mus musculus				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
REFERENCE	1 Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)				
AUTHORS	9279253				
JOURNAL	MEDLINE				
PUBMED	10349636				
REFERENCE	2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)				
AUTHORS	20493574				
JOURNAL	MEDLINE				
PUBMED	11042159				
REFERENCE	3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer Genome Res. 10 (11), 1757-1771 (2000)				
AUTHORS	20530913				
JOURNAL	MEDLINE				
PUBMED	11076861				
REFERENCE	4 The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)				
AUTHORS	5 The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)				
JOURNAL	MEDLINE				
PUBMED	11076861				
REFERENCE	6 (bases 1 to 1864) Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akaira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)				
AUTHORS	TITLE				
JOURNAL	JOURNAL				
PUBMED	COMMENT				
RESULT 12	AK053497				

307	Db	GGCTAACGTTGCCATCTCTGTCTCATCTCCAGTTTGAGGTGTCTATGTCTCAGCTGGTGAA	366
183	QY	CGTGGACGAGAAATCAACTACTTATAACCAATATATGCTGTCTGTTGAGTGGAAATGA	242
367	Db	GGTGGATGAAGTAAACCGATCATGGAAACCACTGTGGCTGAAGCAAAATCTGGNAATGA	426
243	QY	CTACAACCTGAGGTGGAACGACGCGAGTATGCGGGGTCAAGGACCTCAGGATCACGCC	302
427	Db	CTACAACCTGAAATGGAAACCCCTCTGACTACCAAGGGTGGAGTTTCATGCGAGTCCCTGC	486
303	QY	CAACAAGTTGTGGAGCGGACGCTCTTATGTATATAGTCTGACGAGGGTTTTCACGG	362
487	Db	AGAGAAGATCTGGAAACACAGACATCGTGTCTTAAACAACGCCGATGGGGATTTCCAAGT	546
363	QY	GACCTACGACCAACGTTGTTGTCAGAACGCGCGCAGTTGCTGTACGTGCACCTCG	422
547	Db	GGATGACAAACCAAAGCTCTACTCAAGTACACAGAGAAGTACATTGGATCCCTCCGCG	606
423	QY	CATATTTCAAGACGACATGCAAGATGGACATCGGGTGGTTTCCCTTCGACGACCAACATGT	482
607	Db	CATCTTTAAGAGCTCATGCAAAATCGATGTGACCTCTCCCGCTTTGACTACCCAAAACGTG	666
483	QY	TGATATGAAGTTTCGGTAGCTGGACATATGACGGCAATCAGTTGGATCTGCTGCTAAAAAG	542
667	Db	CACCATGAAGTTTCGGTCCCTGGTCTCTGCAAGGCAAGATCGACTGGTCTCTCA---T	722
543	QY	TGAGGACGGCGCGATCTATCGGACTTTCATAACAATGGGAGTGGTATCTAATAGGAAT	602
724	Db	TGGCTCTTCAATGAACCTCAAGGACTATTGGGAAAGTGGCAGTGGGCCATCATTAAGC	783
603	QY	GCCAGGCAAAAAGAACCAATAACATACGGTGTGCCCCGAGCCCTPAGGTGACGTCAC	662
784	Db	CCGGGCTACAAACATGAATCAAGTACAACTGTGTGAGGAGATCTACCAAGACATCAC	843
663	QY	CTTCAACATCATGATGAAGAAGCAACCTTGTTACTCTTCTTCAACTGATCTCCGGTG	722
844	Db	GTACTTCGTATACATTCGCCGCCCTGCCGCTGTTCTACACCATCAACCTCATCTCCG	903
723	QY	CGTCTGTATCTCATGATGGACATCTTCGGCTTCACACTGCCACGACTCCGAGAGAA	782
904	Db	CTGTCTCATCTCTTCTCTCATCTGTTGCTGTTCTTCTACCTGCCCTCCGACTGTGGGAGAA	963
783	QY	ATCTACACTTGGAGTCACTATTCTTCTATCGCTCAGCGTGTCTCTAACCTGGTAGCCGA	842
964	Db	GGTGACGCTCTGCATCTCCGTGCTCTCTCCCTCAGCGTCTTCTCTCTCGTGTATCACCGA	1023
843	QY	GAACCTGCCACAGGTTCCGACGCTATCCCCCTGTTAGGACGTTACTTCAATTGTCATCAT	902
1024	Db	GACCATCCCTTCCACTCTGTCATCCCCCTGATCGGGAGTACCTCTCTCTCTACTAT	1083
903	QY	GTTTCATGTTAGGTCGTCTGTGGTACTGACTGTGGTACTCAATTTACACCAATCGAAC	962
1084	Db	GAATTTTGTACCTTGTCCATCGTCATCACAGTCTTTGTGCTCAAGCTGACATCAGAAC	1143
963	QY	AGCTGATATACATGAATGCCACAGTGGATAAATTCAGTATTCCTCAATGGTTGCCATG	1022
1144	Db	TCCGACCACACACAGATGCCACTTGGTCAAGGCTGTGTTTTTGAACCTTCTCCCCAG	1203
1023	QY	GATCTGCGAATGTGAGGCCA	1044
1204	Db	GGTCATGTTTATGACTAGGCCA	1225

RESULT 15	AY402876	1442 bp	DNA	linear	GSS 15-DEC-2003
LOCUS	AY402876				
DEFINITION	Homo sapiens CHRN4 Gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.				
ACCESSION	AY402876				
VERSION	AY402876.1	GI:39758859			
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
1 (bases 1 to 1442)	
REFERENCE AUTHORS	Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL	Science 302 (5652), 1960-1963 (2003)
PUBMED	14671302
2 (bases 1 to 1442)	
REFERENCE AUTHORS	Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.
TITLE	Direct Submission
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES	Location/Qualifiers
source	1. .1442 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" <1. .>.1442 /gene="CHRNA4" /locus_tag="HCML370"
gene	
ORIGIN	

```
Db 599 -----CTACGTGGACGTGACTTAGCACTTCATCATCAAGCGAAGCCTCTGTCTTAC 650
QY 700 TTCTTTCAACCTGATCGTCCGGTCGGTGGTGAATCTCATTCGATGGCACTCTCGGCTTCACA 759
Db 651 ACCATCAACCTCATCATCCCTCGTGGTCTCACCACCTTGTGGCCATCTCGTCTTCTAC 710
QY 760 CTGCCACCGACTCCGGAGAGAACTCACACTTGGAGTCACTATTCTTCTATCGCTGAG 819
Db 711 CTGCCATCCGACTCGCGGAGAGATGACACTGTGCATCTCAGTGTCTGGCACTGACA 770
QY 820 GTGTTCTCAACCTGGTAGCCGAGACCCTGCCACAGGTCTCCGACGCTATCCCCCTGTTA 879
Db 771 TTCTTCTGCTGCTCATCTCTCAAGATCGTGGCCACCCACCTCCCTCGATGTGCTCTCATC 830
QY 880 GGGACGTACTTCAATTGCATCATGTTCAATGGTAGCGTCTGTGGTACTGACTGTGGTG 939
Db 831 GGCAAGTACCTCATGTTCACCATGGTGTGGTCACTTCTTCATCGTCAACGCGTCTGT 890
QY 940 GTACTCAATTACCACCATCGAACAGCTGATATATATGAAATGCCACAGTGGATAAATCA 999
Db 891 GTGCTCAATGTGCACCAACCGCTCGCCAGCACCACACCATGGCACCTGGGTCAAGCGC 950
QY 1000 GTATTCCTCAATGGTTGCCATGGATGCTGCGAATGTCCGAGGCCAGGGAAGAGATCAC 1059
Db 951 TGCTTCTGCAACAAGCTGCTACCTTCTTCTTCAAGCGCCCTGGCCCCGACAGCAGC 1010
```

Search completed: May 8, 2004, 06:38:57
Job time : 4044.19 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:09:28 ; Search time 43.0968 Seconds
(without alignments)

3284.615 Million cell updates/sec

Title: US-09-303-232-6

Perfect score: 2640

Sequence: 1 MAPMLAALALLPVSEOG.....LFTIATVAVLLSAPHIIVQ 501

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: geneseqp1980s:*

2: geneseqp1990s:*

3: geneseqp2000s:*

4: geneseqp2001s:*

5: geneseqp2002s:*

6: geneseqp2003as:*

7: geneseqp2003bs:*

8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2640	100.0	501	3 AAY50816	Aay50816 H. viresc
2	1803.5	68.3	496	3 AAY50815	Aay50815 H. viresc
3	1609	60.9	770	3 AAY50814	Aay50814 D. melano
4	1570.5	59.5	498	4 ABB60432	Abb60432 Drosophil
5	1283	48.6	311	4 ABB63683	Abb63683 Drosophil
6	1258.5	47.7	502	2 AAW44153	Aaw44153 Human neu
7	1258.5	47.7	502	2 AAW09025	Aaw09025 Neuronal
8	1258.5	47.7	502	3 AAB24088	Aab24088 Human PRO
9	1258.5	47.7	502	4 AAB82690	Aab82690 Nicotinic
10	1258.5	47.7	502	4 AAB82690	Aab82690 Nicotinic
11	1258.5	47.7	502	5 ABB70492	Abb70492 Human neu
12	1258.5	47.7	502	5 ABB82435	Abb82435 Human neu
13	1258.5	47.7	502	7 ADA10874	Ada10874 Human neu
14	1258.5	47.7	502	7 ADA47051	Ada47051 Human PRO
15	1258.5	47.7	502	7 ADE57310	Ade57310 Human PRO
16	1254.5	47.5	502	2 AAW69216	Aaw69216 V274T var
17	1253	47.5	502	2 AAW12368	Aaw12368 Neuronal
18	1252.5	47.4	502	4 AAB50015	Aab50015 Mutant hu
19	1248.5	47.3	502	4 AAB50016	Aab50016 Mutant hu
20	1246.5	47.2	502	7 AAB57308	Aab57308 Rat Prote
21	1246.5	47.2	502	7 ADA47049	Ada47049 Rat Prote
22	1242.5	47.1	502	4 AAB50017	Aab50017 Mutant hu
23	1226.5	46.5	511	2 AAW12369	Aaw12369 Neuronal
24	1080	40.9	554	4 AAE12824	Aae12824 Caenorhab
25	1080	40.9	554	6 ABP96318	Abp96318 Caenorhab

ALIGNMENTS

RESULT 1

AAy50816
ID AAY50816 standard; protein; 501 AA.

XX AC AAY50816;

XX DT 17-FEB-2000 (first entry)

XX DE H. virescens acetyl-choline receptor protein from clone Hva7-2.

XX KW Acetyl-choline receptor; nicotinic; insect; insecticide; screening;
XX NW neurotransmission; plant protection agent; conductance; AChR.

XX OS Heliothis virescens.

XX PN DE19819829-A1.

XX PD 11-NOV-1999.

XX PF 04-MAY-1998; 98DE-01019829.

XX PR 04-MAY-1998; 98DE-01019829.

XX (FARB) BAYER AG.

XX PI Adamczewski M, Oellers N, Schulte T;

XX DR WPI; 2000-014207/02.

XX DR N-PSDB; AAZ24477.

XX PT New nucleic acid encoding a nicotinic acetylcholine receptor from insects, used to identify potential insecticides.

XX Example 1a; Page 22-23; 26pp; German.

XX This invention describes a novel nucleic acid (NA) encoding a nicotinic acetyl-choline receptor (I) from insects which can be used as an insecticide. Inhibitors of (I) interfere with neurotransmission. (I) (also vectors containing it, its regulatory regions, and antibodies directed against (I)-encoded proteins) are used to screen for: (a) plant protection agents that alter conductance of AChR, potentially useful as insecticides, or (b) genes which encode polypeptides that are involved in formation of functionally related AChR in insects. (I) are also used to isolate and characterize the specified regulatory regions and for recombinant production of (II). This sequence represents an acetyl-choline receptor isolated from *Heliothis virescens*

XX SQ Sequence 501 AA;

Ade57314 Human Pro
Ade57318 Human Pro
Aaw09022 Neuronal
Ada10857 Human neu
Ada83810 Human CHR
Ade57312 Rat Prote
Ade57316 Rat Prote
Aab50014 ChimERIC
Abb82431 Human neu
Abb61954 Drosophil
Aab50018 Mature ce
Aaw44156 Human neu
Aaw09018 Neuronal
Abb70491 Human neu
Abb82434 Human neu
Ada10863 Human neu
Abb62727 Drosophil
Aaw44155 Human neu
Aar73966 Alpha 2 s

26 989.5 37.5 503 7 ADE57314
27 989.5 37.5 503 7 ADE57318
28 989.5 37.5 504 2 AAW09022
29 989.5 37.5 504 7 ADA10857
30 982.5 37.2 502 6 ADA83810
31 977.5 37.0 499 7 ADE57312
32 977.5 37.0 499 7 ADE57316
33 962.5 36.5 470 4 AAB50014
34 946.5 35.9 504 5 ABB70488
35 946.5 35.9 504 5 ABB82431
36 943 35.7 576 4 ABB61954
37 942 35.7 448 4 AAB50018
38 940.5 35.6 504 2 AAW44156
39 933 35.3 494 2 AAW09018
40 933 35.3 494 5 ABB70491
41 933 35.3 494 5 ABB82434
42 933 35.3 494 7 ADA10863
43 933 35.3 580 4 ABB62727
44 922.5 34.9 529 2 AAW44155
45 922.5 34.9 529 2 AAR73966

Query Match 100.0%; Score 2640; DB 3; Length 501;
 Best Local Similarity 100.0%; Pred. No. 1.2e-267;
 Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPMLAALALLPVSEQGPHEKRLINALLANVTLEPVAESEPLEVRFGLTLQOII 60
 DB 1 MAPMLAALALLPVSEQGPHEKRLINALLANVTLEPVAESEPLEVRFGLTLQOII 60

QY 61 DVEKNQLITNWLSEWNYLRNDSYGGVKDLRITPNKLVKPDVLMYNSADSGFD 120
 DB 61 DVEKNQLITNWLSEWNYLRNDSYGGVKDLRITPNKLVKPDVLMYNSADSGFD 120

QY 121 GTYQTNVVRSGGSLVPPGIFKSTCKMDIAFPDDQHCMDKFGSWTYDGNQLDLVK 180
 DB 121 GTYQTNVVRSGGSLVPPGIFKSTCKMDIAFPDDQHCMDKFGSWTYDGNQLDLVK 180

QY 181 DEAGGDSLDFITNGEWYLIQMPGKNTITYACCEPYVDVTFITMIRRTLYFFNLIVP 240
 DB 181 DEAGGDSLDFITNGEWYLIQMPGKNTITYACCEPYVDVTFITMIRRTLYFFNLIVP 240

QY 241 CVLISSMALLGFTLPDPSGKLTGLVTLISLTVFLNLVAETLPQVSDAIPLLGTYPNCI 300
 DB 241 CVLISSMALLGFTLPDPSGKLTGLVTLISLTVFLNLVAETLPQVSDAIPLLGTYPNCI 300

QY 301 MFVASSVLTVVVLYNHHRTADIHEMPOWIKSVFLQWLPWILRMSRPGKTKITIMN 360
 DB 301 MFVASSVLTVVVLYNHHRTADIHEMPOWIKSVFLQWLPWILRMSRPGKTKITIMN 360

QY 361 TRMRELEKRSKSLANVLDIDDDFRHGGPPPNSTASTGNLPGGCSIFRTDPRRSFVR 420
 DB 361 TRMRELEKRSKSLANVLDIDDDFRHGGPPPNSTASTGNLPGGCSIFRTDPRRSFVR 420

QY 421 PSTMEDVGGGLGSHHRLHLILRELQITARMKKADEAEELISDWKFAAMVDRFCLFV 480
 DB 421 PSTMEDVGGGLGSHHRLHLILRELQITARMKKADEAEELISDWKFAAMVDRFCLFV 480

QY 481 TLFITIAVALLSAPHIIQ 501
 DB 481 TLFITIAVALLSAPHIIQ 501

RESULT 2
 AAY50815
 ID AAY50815 standard; protein; 496 AA.
 AC AAY50815;
 DT 17-FEB-2000 (first entry)
 XX H. virescens acetyl-choline receptor protein from clone Hva7-1.
 DE Acetyl-choline receptor; nicotinic; insect; insecticide; screening;
 XX neurotransmission; plant protection agent; conductance; ACHR.
 OS Heliothis virescens.
 PN DE19819829-Al.
 PD 11-NOV-1999.
 XX 04-MAY-1998; 98DE-01019829.
 XX 04-MAY-1998; 98DE-01019829.
 XX (FARB) BAYER AG.
 XX Adamczewski M, Oellers N, Schulte T;
 XX WPI; 2000-014207/02.
 XX N-PSDB; AAZ24476.
 PT New nucleic acid encoding a nicotinic acetylcholine receptor from

PT insects, used to identify potential insecticides.
 XX Example 1a; Page 17-19; 26pp; German.
 XX This invention describes a novel nucleic acid (NA) encoding a nicotinic
 CC acetyl-choline receptor (I) from insects which can be used as an
 CC insecticide. Inhibitors of (I) interfere with neurotransmission. (I)
 CC (also vectors containing it, its regulatory regions, and antibodies
 CC directed against (I)-encoded proteins) are used to screen for: (a) plant
 CC protection agents that alter conductance of ACHR, potentially useful as
 CC insecticides, or (b) genes which encode polypeptides that are involved in
 CC formation of functionally related ACHR in insects. (I) are also used to
 CC isolate and characterize the specified regulatory regions and for
 CC recombinant production of (II). This sequence represents an acetyl-
 CC choline receptor isolated from *Heliothis virescens*
 XX Sequence 496 AA;

Query Match 68.3%; Score 1803.5; DB 3; Length 496;
 Best Local Similarity 68.7%; Pred. No. 8.2e-180;
 Matches 347; Conservative 53; Mismatches 78; Indels 27; Gaps 7;

QY 2 APMLAALALLPVSEQGPHEKRLINALLANVTLEPVAESEPLEVRFGLTLQOII 60
 DB 12 APAGLLLLCLLWPRGARCGYHEKRLIHLLDHYNVLERPVWVESDPLQLSFGLTLQOII 71

QY 61 DVEKNQLITNWLSEWNYLRNDSYGGVKDLRITPNKLVKPDVLMYNSADSGFD 120
 DB 72 DVEKNQLITNWLSEWNYLRNDSYGGVKDLRITPNKLVKPDVLMYNSADSGFD 131

QY 121 GTYQTNVVRSGGSLVPPGIFKSTCKMDIAFPDDQHCMDKFGSWTYDGNQLDLVK 180
 DB 132 STYPTNVVRNNGSLVPPGIFKSTCKIDITWFPDDQRCMKFGSWTYDGYQLDLQ 191

QY 181 DEAGGDSLDFITNGEWYLIQMPGKNTITYACCEPYVDVTFITMIRRTLYFFNLIVP 240
 DB 192 DEGGGDISFVINGEWELIGVPGKRNEIYVNCCEPYDITFAVWIRRTLYFFNLIVP 251

QY 241 CVLISSMALLGFTLPDPSGKLTGLVTLISLTVFLNLVAETLPQVSDAIPLLGTYPNCI 300
 DB 252 CVLISSMALLGFTLPDPSGKLTGLVTLISLTVFLNLVAETLPQVSDAIPLLGTYPNCI 311

QY 301 MFVASSVLTVVVLYNHHRTADIHEMPOWIKSVFLQWLPWILRMSRPGKTKITIMN 360
 DB 312 MFVASSVLTVVVLYNHHRTADIHEMPOWIKSVFLQWLPWILRMSRPGSATPPARV- 370

QY 361 TRMRELEKRSKSLANVLDIDDDFRHGGPPPNSTASTGNLPGGCSIFRTDPRRSFVR 420
 DB 371 PPPDLELRERSKSLANVLDIDDDFRH- - - - -PQAQQ- - - - -YRGG- - - 414

QY 421 PSTMEDVGGGLGSH- - - - -HRELHLILRELQITARMKKADEAEELISDWKFAAMVDRF 475
 DB 415 - - - - -EENGAGLAHSCFVDYELSLILKEIRVITDQMKRDEDEDADISRDWKAAMVDR 470

QY 476 CLFVFTLTITIAVALLSAPHIIQ 500
 DB 471 CLFIITLTITIAVALLSAPHIMV 495

RESULT 3
 AAY50814
 ID AAY50814 standard; protein; 770 AA.
 XX AAY50814;
 AC AAY50814;
 XX 17-FEB-2000 (first entry)
 DT D. melanogaster acetyl-choline receptor protein from clone Da7.
 DE Acetyl-choline receptor; nicotinic; insect; insecticide; screening;
 XX neurotransmission; plant protection agent; conductance; ACHR.
 KW Drosophila melanogaster.
 XX

XX DE19819829-A1.
PN 11-NOV-1999.
PD 04-MAY-1998; 98DE-01019829.
PF 04-MAY-1998; 98DE-01019829.
PR (FARB) BAYER AG.
PA Adamczewski M, Oellers M, Schulte T;
XX WPI: 2000-014207/02.
XX N-PSDB; AAZ24475.
XX New nucleic acid encoding a nicotinic acetylcholine receptor from
PT insects, used to identify potential insecticides.
PS Example 1a; Page 12-14; 26pp; German.
XX This invention describes a novel nucleic acid (NA) encoding a nicotinic
CC acetyl-choline receptor (I) from insects which can be used as an
CC insecticide. Inhibitors of (I) interfere with neurotransmission. (II)
CC (also vectors containing it, its regulatory regions, and antibodies
CC directed against (I)-encoded proteins) are used to screen for: (a) plant
CC protection agents that alter conductance of AChR, potentially useful as
CC insecticides, or (b) genes which encode polypeptides that are involved in
CC formation of functionally related AChR in insects. (I) are also used to
CC isolate and characterize the specified regulatory regions and for
CC recombinant production of (II). This sequence represents an acetyl-
CC choline receptor isolated from *Drosophila melanogaster*
XX
XX Sequence 770 AA;
Query Match 60.9%; Score 1609; DB 3; Length 770;
Best Local Similarity 60.8%; Pred. No. 4.2e-159;
Matches 319; Conservative 57; Mismatches 65; Indels 84; Gaps 10;
QY 10 LLALLPVSEQ-----GHEKRLNALLANTYLERPVANSEPLEVRFGLTQIIVDEK 65
DB 295 LITYLNSAKVCLAGYHEKRLLDLPYNTLERPVNESDPLQLSFGTLTQIIVDEK 354
QY 66 NOLLITNLSLEWYNDYLRWDSYGVKDLRITPNKWKDPVLMYNSADEFGDTYOT 125
DB 355 NQLLVTVNWLKLEWNNLAWNTSDTGGVKDLRIPHRWKDPVLMYNSADEFGDTYOT 414
QY 126 NVVVRSGSCLYVPPGIFKSTCKMDIAWFPDDQHCMDKFGSWTYDGNQLDLVKDEAG 185
DB 415 NVVVRNNGSCLYVPPGIFKSTCKIDITWFPDDQRCMEKFGSWTYDGFQLDLQDETGG 474
QY 186 DLSDFITNGEWLIGMPGKKNITYIACCPBPYVDVTFIMIRRTLYFFNLIIVCVLIS 245
DB 475 DISSVYLVNGEWELLVGPGKRNELIYNCCPEYIDITFAIIRRTLYFFNLIIPCVLIA 534
QY 246 SMALLGFTLPPSGEKLTLGVITLLSLTVFLNVAETLPQVSDAIPLLGTFFNCFMFWA 305
DB 535 SMALLGFTLPPSGEKLTLGVITLLSLTVFLNVAETMPATSDAVPL----- 581
QY 306 SSVVLTVVLYNHYHRTADIHMPQWTKSVFLOLWPLWILMRSPGKKITRK--TIMMNT 362
DB 582 -----WIRIVLCWLPWILMRSPGRPLILEFFPTPCSDTS 617
QY 363 -----MRELEKERSKSLANVLIDDDDFRHG--PPPNSTAGTGMGPGCSIFRTD 413
DB 618 SERKHQILSDVELKERSKSLANVLIDDDDFRHNCRPMTG-----GTLPHNPAFYRTV 672
QY 414 FRRSFVRPSTMDVG--GGIGS-----HHR-----ELHLTLRELQRTARMKA 455
DB 673 YCGG-----DDGIGIPGSTRMPDVAHTTCIKSSTSEYELGLILKEIRFITDQLRKD 724
QY 456 DEEAELISDMKFAAMVVDVRFCLVFVFTLTIIATVAVLLSAPHIIV 500
|| : : : ||||| || : : : ||||| || : : : ||||| || : : : ||||| |||||

DB 725 DECNDIANDWKFAAMVVDRLCLIFITMFAIATIAVLISAPHIIV 769
RESULT 4
ID ABB60432
XX ABB60432 standard; protein; 498 AA.
AC ABB60432;
XX 26-MAR-2002 (first entry)
XX *Drosophila melanogaster* polypeptide SEQ ID NO 8088.
DE *Drosophila*; developmental biology; cell signalling; insecticide;
XX *Drosophila*; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX *Drosophila melanogaster*.
OS WO200171042-A2.
XX
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US009231.
XX 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX (PEKE) PE CORP NY.
XX Venter JC, Adams M, Li PWD, Myers EW;
PI WPI: 2001-656860/75.
XX N-PSDB; ABL04535.
DR
XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from *Drosophila* and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 8088; 2lpp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 498 AA;
Query Match 59.5%; Score 1570.5; DB 4; Length 498;
Best Local Similarity 61.1%; Pred. No. 2.4e-155;
Matches 319; Conservative 44; Mismatches 84; Indels 75; Gaps 9;
QY 2 APMLAAL-----LALLPVSEQGPHEKRLNALLANTYLERPVANSEPLEVRFGLT 55
DB 3 SPLPASLSFLVLLIFLAITKESQGPHEKRLNALLSTYNTLERPVANSEPLEVRFGLT 62
QY 56 LQIIDVDEKQOLLITNLSLEWYNDYLRWDSYGVKDLRITPNKWKDPVLMYNSA 115
DB 63 LQIIDVDEKQOLLITNLSLEWYNDYLRWDSYGVKDLRITPNKWKDPVLMYNSA 122
QY 116 DSGFGDTGYTNVVRSGSCLYVPPGIFKSTCKMDIAWFPDDQHCMDKFGSWTYDGNQL 175
DB 123 DSGFGDTGYTNVVRSGSCLYVPPGIFKSTCKIDITWFPDDQRCMEKFGSWTYDGNQL 182
QY 176 DLVLKDEAGDLSDFITNGEWLIGM-PGKKNITYIACCP-----EPYVDVT 221
DB 183 DLVLNSEDGDLSDFITNGEWLIGVHAGKEBYDSLRLLPRTICRYHLVYTNSSPVI--- 239
|| : : : ||||| || : : : ||||| || : : : ||||| || : : : ||||| |||||

```
QY 222 FTIMIRRTLYFFNLIIVPCVLSSMALL-----GFTLPDPSGEKLTILGV----- 266
Db 240 -----ILFFQNCAMCANLIDGPTGLHIAAGFGRETDAGNYTTIINSIKPCRVRH 291
QY 267 -----TILLSLTVFLNLVAETLPQVSDAIPLLGTGFYFNCIMFMWASSVVLTV 312
Db 292 ADVNGCCSSYRYTHTLKRYSLKYGPPIGRITDIHYQC-RAESTVFCIMFMWASSVVLTV 350
QY 313 VVLNYHRTADIHEMPQWIKSVFLOLWPILWMSRPGKKTIRKTIIMNTMRELELKERS 372
Db 351 VVLNYHRTADIHEMPQWIKSVFLOLWPILWMSRPGKKTIRKTIILSNMKELELKERS 410
QY 373 SKSLLANVLDDIDDFRHGPPPPNSTAGTGNLGPCCSIFRTDFRSFVRPSTMEDVGGGLG 432
Db 411 SKSLLANVLDDIDDFRH-----TISGQTAIGSS-----ASFGRPITVEHHTAIG 456
QY 433 SHRELHLILREIQFITARMKKADDEAEELISDMKFAAMVVDR 474
Db 457 CNEKDLHLILKELQFITARMKKADDEAEELIGDMKFAAMVVDR 498

RESULT 5
ABB63683
ID ABB63683 standard; protein; 311 AA.
XX
AC ABB63683;
XX
DT 26-MAR-2002 (first entry)
XX
KW Drosophila melanogaster polypeptide SEQ ID NO 17841.
XX
DE Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
DR N-PSDB; ABL07786.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Disclosure; SEQ ID NO 17841; 2lpp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS7737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 311 AA;
```

Query Match 48.6%; Score 1283; DB 4; Length 311;
Best Local Similarity 79.4%; Pred. No. 1.8e-125;

```
Matches 231; Conservative 30; Mismatches 30; Indels 0; Gaps 0;
QY 4 MLAALALLLPVSEQGPHEKRLINALLANYTLERPVANESPELVRFGLTQQIITDVD 63
Db 20 LVTGLGLINIPACAAQPHKRLHLLDNDYNSLERPVVNESDPLQLSFGTLTQQIITDVD 79
QY 64 EKNQLLITNLWLSLEWVDNLRWDSYGGVKDLRITPNKLWKPDLVLMYNSADEGFGDTY 123
Db 80 EKNQLLITNLWLSLEWVDNLRWDSYGGVKDLRITPNKLWKPDLVLMYNSADEGFGDTY 139
QY 124 QTVVVRSGSCLYVPPGIFKSTCKMDIAWFPDDQHCMDKFGSWTYDGNQLDLVLKDEA 183
Db 140 ATNVVNRNGSCLYVPPGIFKSTCKDITWFPDDQRCCKMFGSWTYDGFQLDLQLEA 199
QY 184 GGLSDFITNGEWWLGMPOKNTIYACCPYVVDVTFIMIRRTLYFFNLIIVPCVL 243
Db 200 GSDISSFITNGEWWLGMPOKNTIYACCPYVVDVTFIMIRRTLYFFNLIIVPCVL 259
QY 244 ISSMALLGFTLPDPSGEKLTILGVTTILSLTVFLNLVAETLPQVSDAIPLLG 294
Db 260 IASMALLGFTLPDPSGEKLTILGVTTILSLTVFLNLVAETLPQVSDAIPLLG 310

RESULT 6
AAW44153
ID AAW44153 standard; protein; 502 AA.
XX
AC AAW44153;
XX
DT 14-MAY-1998 (first entry)
XX
DE Human neuronal nicotinic acetylcholine receptor alpha-7 subunit.
XX
KW Human; neuronal nicotinic acetylcholine receptor; alpha-7 subunit;
KW brain tissue; screening; NACHR; antibody.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT Peptide 1..23
FT Domain /label= signal
FT Domain 229..256
FT Domain /label= TMD1
FT Domain /note= "transmembrane domain"
FT Domain 262..284
FT Domain /label= TMD2
FT Domain /note= "transmembrane domain"
FT Domain 290..317
FT Domain /label= TMD3
FT Domain /note= "transmembrane domain"
FT Domain 318..461
FT Domain /label= cytoplasmic_loop
FT Misc-difference 343 /note= "encoded by CTS"
FT Domain 462..487
FT Domain /label= TMD4
FT Domain /note= "transmembrane domain"
XX
PN WO9420617-A2.
XX
PD 15-SEP-1994.
XX
PF 08-MAR-1994; 94WO-US002447.
XX
PR 08-MAR-1993; 93US-00028031.
XX
PA (SALK ) SALK INST BIOTECHNOLOGY IND ASSOC.
XX
PI Elliott KJ, Ellis SB, Harpold MM;
XX
DR WPI; 1994-303024/37.
DR N-PSDB; AAV12197.
XX
```


Db 461 KFAACVVDRLCLMAFSVFTIICITIGILMSAPNFV 494

RESULT 8

AAB24088
ID AAB24088 standard; protein; 502 AA.

XX AC AAB24088;

XX 29-JAN-2001 (first entry)
XX Human PRO2145 protein sequence SEQ ID NO:77.

XX Human; tumour; diagnosis; neoplastic disease; neoplastic cell growth;
KW proliferation; tumorigenesis; identification; cancer; cytostatic;
KW neurotropic; neuroprotective; antiinflammatory; immunosuppressive;
KW immunostimulant; antiangiogenic; leukaemia; lymphoid malignancy;
KW neuronal disorder; glial disorder; astrocytal disorder; angiogenic;
KW hypothalamic disorder; glandular disorder; macrophagal disorder;
KW epithelial disorder; stromal disorder; blastocoelec disorder;
KW inflammatory disorder; immunologic disorder.

XX Homo sapiens.

OS
XX WO200053755-A2.

XX 14-SEP-2000.

XX 06-JAN-2000; 2000WO-US000376.

XX 08-MAR-1999; 99WO-US005028.

XX 02-JUN-1999; 99WO-US012252.

XX 23-JUN-1999; 99US-0141037P.

XX 07-JUL-1999; 99US-0143048P.

XX 26-JUL-1999; 99US-0145698P.

XX 30-NOV-1999; 99WO-US028313.

XX 20-DEC-1999; 99WO-US030911.

XX 05-JAN-2000; 2000WO-US000219.

XX (GETH) GENENTECH INC.

XX Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hillan KJ, Roy MA;
XX Watanabe CK, Wood WI;
XX WPI; 2000-572270/53.

XX N-PSDB; AAC58395.

XX Thirty PRO polynucleotides encoding PRO polypeptides, useful in the

XX treatment, diagnosis and prevention of cancer.

XX Claim 61; Fig 58; 286pp; English.

XX The present invention describes an isolated antibody that binds to one of
XX the human PRO proteins designated PRO212, PRO290, PRO341, PRO355, PRO619,
XX PRO717, PRO809, PRO830, PRO848, PRO943, PRO1005, PRO1009, PRO1025,
XX PRO1030, PRO1097, PRO1107, PRO1111, PRO1153, PRO1182, PRO1184, PRO1187,
XX PRO1281, PRO23, PRO39, PRO834, PRO1317, PRO1710, PRO2094, PRO2145 OR
XX PRO2198. PRO antagonists can be used to inhibit tumour cell growth. The
XX PRO polypeptides and nucleotides are useful in the treatment, diagnosis
XX and prevention of cancer. The antibodies and other anti-tumour compounds
XX maybe used to treat various conditions, including those characterised by
XX overexpression and/or activation of the amplified PRO genes. Exemplary
XX conditions or disorders to be treated with such antibodies and other
XX compounds include benign or malignant tumours (e.g., renal, liver,
XX kidney, bladder, breast, gastric, ovarian, colorectal, prostate,
XX pancreatic, lung, vulva, thyroid, hepatic carcinomas, sarcomas,
XX glioblastomas, and various head and neck tumours), leukaemias and
XX lymphoid malignancies, other disorders such as neuronal, glial,
XX astrocytal, hypothalamic and other glandular, macrophagal, epithelial,
XX stromal and blastocoelec disorders, and inflammatory, angiogenic and
XX immunologic disorders. AAC58242 to AAC58366 represent PCR primers and
XX hybridisation probes used in the isolation of the human PRO sequences.
XX AAC58367 to AAC58396 and AAB24057 to AAB24089 represent human PRO

CC polynucleotide and protein sequences given in the exemplification of the
CC present invention
XX Sequence 502 AA;

Query Match 47.7%; Score 1258.5; DB 3; Length 502;
Best Local Similarity 48.6%; Pred. No. 1.4e-122;
Matches 250; Conservative 75; Mismatches 138; Indels 51; Gaps 7;

Qy 8 LALLA-LLPVSEOGPHEKRLINALLANYTLERPVANSEPLEVRFGITLQOIIVDVEKN 66
Db 10 LALAASLLHVSLOGEFQKLYKELVKYNPLERVANDSQPLTYVYLSLLQIMVDVEKN 69

Qy 67 QLLITNWLSEWNYLNRWDSBYGVKDLRITPNKWKPDVLMVNSADEGFGTQTN 126
Db 70 QVLTNINLQMSWTDHYLQWVSEYPGVTVRPDQGIWKPDILLYNSADRFATHTN 129

Qy 127 VVVRSGSCLYVPGIFKSTCKMDIAWFPDDQDCMKFGSWTYDGNLDLVLKDEAGD 186
Db 130 VLVNSSGHCQYLPPIGFKSSCYIDVRWFPDVQCKLKFGSWSYCGMSLDLQMQE---AD 186

Qy 187 LSDFITNGEWYLLGMPGKNITTYACCPYVDVTFITMIRRTLYYFFNLIVPCVLIS 246
Db 187 ISGYIPNGEWDLVGIPGKRSERFYECKEYPDVTFVTWRRRTLYYGLNLLIECVLISA 246

Qy 247 MALLGFTLPPDSGKLTGVTILLSLTVFLNLVAETLPQVSDAIPILGTFFNCFMFMVAS 306
Db 247 LALLVFLPADSGEKISLGITVLLSVFVLLVAEIMPATSDSVFLTAQYFASTMIIVGL 306

Qy 307 SVLVTVVLYNHHRTADIHEMPQWIKSVFLQWLPWILMRSPGKKITRKTIMMTRREL 366
Db 307 SVVTVIVLQVHHHDPDGKMKWTRVILLNWCWFLMRKRPBGDKVRPAQCHKQRCSL 366

Qy 367 ELKERSKSLANVLDDDDFRHGGPPPNSTASGNL-----GGCC 407
Db 367 ASVEMSAVA-----PPP-----ASGNLLYIGFRGLDGVHCVPTPDSDGVVC 407

Qy 408 SIFRTDFRSFVRPSTMEDV--GGGLGSHHRELHLILRELQFITARMKKADEEAELISDW 465
Db 408 G-----RMACSPTHDEHLLHGGQPEGDPDLAKILLEEVRYIANRRCQDESAVCSW 460

Qy 466 KFAAMVDRPCLFVFTLTFTIATVAVLLSAPHII 499
Db 461 KFAACVVDRLCLMAFSVFTIICITIGILMSAPNFV 494

RESULT 9

AAB82690
ID AAB82690 standard; protein; 502 AA.

XX AC AAB82690;

XX 15-OCT-2001 (first entry)

XX Nicotinic acetylcholine receptor alpha7.

XX Nicotinic acetylcholine receptor; nAChR; human;
KW acetylcholine binding protein; AChBP; mollusc; ligand-binding protein;
KW ligand-gated ion channel; crystal; drug design; protein co-ordinate data;
KW schizophrenia; Alzheimer's disease; nicotine addiction;
KW Tourette's syndrome; therapy; neurotropic; neuroprotective.
XX Homo sapiens.

PH Key Location/Qualifiers
FT Region 108..115

FT /note= "conserved ligand-binding region, residues Trp108
FT and Tyr115 are essential"

FT Region 171..173

FT /note= "conserved ligand-binding region, residues Trp171
FT and Tyr173 are essential"

FT Region 210..217

FT /note= "conserved ligand-binding region, residues Tyr210,

PT Cys212, Cys213 and Tyr217 are essential"

XX WO200158951-A2.

XX 16-AUG-2001.

XX 09-FEB-2001; 2001WO-EP001457.

XX 10-FEB-2000; 2000EP-00200443.

XX 31-OCT-2000; 2000EP-00203810.

XX (TEWE-) STICHTING TECH WEIENSCHAPPEN.

XX Smit AB, Sixma TK;

XX WPI; 2001-497071/54.

XX Water-soluble ligand-binding proteins derived from molluscs and analogs

PT of ligand-gated ion channels, useful in drug screening assay, where the

PT drugs identified can be used in the treatment of Alzheimer's disease or

PT schizophrenia.

XX Disclosure; Page 252-254; 260pp; English.

XX The present sequence is that of the alpha subunit of human nicotinic

CC acetylcholine receptor (nAChR). The sequence includes regions that are

CC conserved throughout the various nAChR alpha subunits and which are

CC essential for ligand binding. The invention relates to water-soluble

CC ligand-binding proteins derived from molluscs, especially acetylcholine-

CC binding proteins (AChBPs) and analogues of ligand-gated ion channels,

CC their crystals, and their use for screening ligands of ligand-gated ion

CC channels. The water-soluble ligand-binding proteins are capable of

CC forming multimers and are amenable to crystallization. The crystal

CC structure of AChBP is provided, and can be used to generate 3D models of

CC the extracellular ligand-binding domain of ligand-gated ion channels and

CC thus for screening of drugs that act on these ion channels. Chimeric

CC proteins are provided that are capable of binding a ligand of a ligand-

CC gated receptor, and comprise at least the amino acids of the AChBP

CC determining solubility of the AChBP, in the same positions as in the

CC AChBP, and also comprising amino acids determining binding to the ligand.

CC In the chimeric proteins, at least the essential amino acids of at least

CC 1 of the conserved regions of an nAChR have been substituted for the

CC corresponding amino acids, and preferably entire stretches have been

CC substituted. New drugs can be developed that selectively intervene in

CC neuronal signalling pathways, especially where the ligand-gated ion

CC channel is the nAChR, and the related disorder is Tourette's syndrome,

CC Alzheimer's disease, addiction to nicotine or schizophrenia

XX Sequence 502 AA;

SQ Query Match 47.7%; Score 1258.5; DB 4; Length 502;

Best Local Similarity 48.6%; Pred. No. 1.4e-122;

Matches 250; Conservative 75; Mismatches 138; Indels 51; Gaps 7;

QY 8 LALLA-ILPVSEOGHEKRLNALLANVTLPVANESEPLEVFGTLQOIIVDDEKN 66

DB 10 LALAASLLHVSLSQGFQRLKLYKELYNPLRPVANDSQPLTVYFSLQLQIMVDDEKN 69

QY 67 QLLITNIWLSLEWNYLNRWNSDEYGVKDLRITPNKWLKPDVLMYNSADEGFGDTYQTN 126

DB 70 QVLTNIWLSQWTHYLOWNVSEYGVKVTVPFDQIWKPDILLYNSADERFDATFTN 129

QY 127 VVVRSGGSLYPPGIFKSTCKMDIAWFPDQHCMDKFGSWTYDGNQLDLVLKDEAGD 186

DB 130 VLNVSSGHQYLPGLPGIFKSSCVIDRVFPFDVQHCXKLFKGSWYGVSLDLQWQE---AD 186

QY 187 LSDFTNGEYWLIGMPGKKNITTYACCPPEYVDVFTIMRRRTLYFFNLIVPCVLISS 246

DB 187 ISGYIPNGEWDLVGIPGKRSEFYECCKEPYDPVFTVTMRRTLYYGLNLLIPCVLISA 246

QY 247 MALLGFTLPDSEKLTGLTGLTLLSLTVFLNLVAETLPQVSDAIPLLGYFNCIMFVAS 306

DB 247 LALLVFLPADSGEKISLTGLTGLTLLSLTVFLNLVAEIMPATSDSVPLIAQYFASTMIIVGL 306

QY 307 SVVLTVVVLNYHRTADIHEMPQWIKSVFLOLFWILMSRPGKKTIRKTIIMNTRMREL 366

DB 307 SVVTVIVLQYHHDPDGGKPKWTRVILLNWCAWFLMRKRPGEDKVRPACQHKQRCSL 366

QY 367 ELKERSKSLANVLDIDDDFRHGPPPPNPGTASTGNL-----GFGC 407

DB 367 ASVEMSAVA-----PPP---ASNGNLLYIGFRLDGVCHVCVPTPDSGVVC 407

QY 408 SIFRTDFRRSFVRPSTMEDV--GGGLGSHHREHLILRELOFTARMKKADEAEELISDW 465

DB 408 G-----RMACSPTHDEHLHGQPPGPDGLAKILEEVRYIANRFRCDSEAVACSEW 460

QY 466 KFAAMVVDRECLFVFTLTFTIATVAVLLSAPHII 499

DB 461 KFAACVVVDRLCLMAFSVFTIITIGILMSAPNFV 494

RESULT 10

AAB50012

ID AAB50012 standard; protein; 502 AA.

XX AAB50012;

XX AC

XX 14-MAR-2001 (first entry)

XX DT

XX DE Wild-type human alpha7 ligand gated ion channel.

XX KW Human; alpha7 nicotinic acetylcholine gated ion channel;

XX KW 5-hydroxytryptamine; 5-HT3; calcium ion conductance.

XX OS Homo sapiens.

XX WO200073431-A2.

XX PD 07-DEC-2000.

XX PF 25-MAY-2000; 2000WO-US011862.

XX PR 27-MAY-1999; 99US-0136174P.

XX PA (PHAA) PHARMACIA & UPJOHN CO.

XX PI Groppi VE, Wolfe ML, Berkenpas MB;

XX WPI; 2001-061524/07.

XX DR N-PSDB; AAC90380.

XX PT Special cell culture medium for treating cells and for inducing mammalian

PT cell lines to conduct calcium ions, comprising specified concentrations

PT of ions of sodium, calcium and potassium at specified pH.

XX Disclosure; Page 61-63; 77pp; English.

XX The present sequence is wild-type human alpha7 nicotinic acetylcholine

CC gated ion channel. The human alpha7 ion channel was used in the

CC construction of an alpha7/5-hydroxytryptamine (5-HT3) chimeric ligand

CC gated ion channel (see AAC90382 and AAB50014). The alpha7/5-HT3 chimeric

CC ion channel can be expressed by recombinant cells in the present

CC invention, resulting in preferential calcium ion conductance by the cells

XX Sequence 502 AA;

Query Match 47.7%; Score 1258.5; DB 4; Length 502;

Best Local Similarity 48.6%; Pred. No. 1.4e-122;

Matches 250; Conservative 75; Mismatches 138; Indels 51; Gaps 7;

QY 8 LALLA-ILPVSEOGHEKRLNALLANVTLPVANESEPLEVFGTLQOIIVDDEKN 66

DB 10 LALAASLLHVSLSQGFQRLKLYKELYNPLRPVANDSQPLTVYFSLQLQIMVDDEKN 69

QY 67 QLLITNIWLSLEWNYLNRWNSDEYGVKDLRITPNKWLKPDVLMYNSADEGFGDTYQTN 126

DB 70 QVLTNIWLSQWTHYLOWNVSEYGVKVTVPFDQIWKPDILLYNSADERFDATFTN 129

QY 127 VVVRSGGSLYPPGIFKSTCKMDIAWFPDQHCMDKFGSWTYDGNQLDLVLKDEAGD 186

DB 130 VLNVSSGHQYLPGLPGIFKSSCVIDRVFPFDVQHCXKLFKGSWYGVSLDLQWQE---AD 186

QY 187 LSDFTNGEYWLIGMPGKKNITTYACCPPEYVDVFTIMRRRTLYFFNLIVPCVLISS 246

DB 187 ISGYIPNGEWDLVGIPGKRSEFYECCKEPYDPVFTVTMRRTLYYGLNLLIPCVLISA 246

QY 247 MALLGFTLPDSEKLTGLTGLTLLSLTVFLNLVAETLPQVSDAIPLLGYFNCIMFVAS 306

DB 247 LALLVFLPADSGEKISLTGLTGLTLLSLTVFLNLVAEIMPATSDSVPLIAQYFASTMIIVGL 306

Db 70 QVLTNINLQMSWTDHYLQNNVSEYGVKTVRPDQGIWKPDILLNYSADERDFATHTN 129
 QY 127 VVVRSGSCLYVPPGIFKSTCKMDIAWPFDDQHCMDKFGSWTYDGNQLDLVKDEAGD 186
 Db 130 VLVNSSGHCOYLPPIGIFKSCYDVRWFFDVQHCCKLFGSWSYGGSLDLQOE---AD 186
 QY 187 LSDFTNGEWYLGMPGKKNITTYACCPPEYVDVFTTIRRTLYFFNLIVPCVLISS 246
 Db 187 ISGYIPNGEWDLVGIPGKSERFYECKEYFPDVTFTVMRRRTLYYGLNLLIPCVLISA 246
 QY 247 MALLGFTLPDPSGEKLTGLVTLLSLTVFLNVAETLPQVSDAIPLGTYFNCIMFVNAS 306
 Db 247 LALLVFLPADSGEKISLIGITVLLSLTVFLLVAEIMPATSDSVPLIAQYFASTMTIIVGL 306
 QY 307 SVVLTVVVLYNHHRTADIHEMPQWIKSVFLOLPMILRMSRPGKTKITRKTIMMTRMREL 366
 Db 307 SVVTVIVLVYHHHDPDGGKMPKWTIRVILLNWCWFLMRKPGEDKVRPACQHKQRCSL 366
 QY 367 ELKERSKSLLANVLIDDDFRHGPPPNSTAGNL-----GPGC 407
 Db 367 ASVMSAVA-----PPP---ASNGNLLYIGRGLDGVHCVPTPDGSGVVC 407
 QY 408 SIFRTDFRSFVRPSTMEDV---GGGLGSHHRELHLILRELQFTARMKKADEAEELISDW 465
 Db 408 G-----RMACSPHDEHLHGQPPGDPDLAKILEEVRYIANFRCODESEAVCSEW 460
 QY 466 KFAAMVVDVDRFCLFVFTFTIATVAVLLSAPHII 499
 Db 461 KFAACVVDRLCLMAFSVFTIICIGILMSAPNFV 494
 RESULT 11
 ID ABG70492 standard; protein; 502 AA.
 AC ABG70492;
 DT 06-DEC-2002 (first entry)
 DE Human neuronal nicotinic acetylcholine receptor alpha 7 subunit.
 KW Human; neuronal nicotinic acetylcholine receptor; nNACHr; receptor;
 KW ion flux; alpha 7 subunit.
 OS Homo sapiens.
 PN US6440681-B1.
 PD 27-AUG-2002.
 PF 07-JUN-1995; 95US-00487596.
 PR 03-APR-1990; 90US-00504455.
 PR 30-NOV-1992; 92US-00938154.
 PR 08-MAR-1993; 93US-00028031.
 PR 08-NOV-1993; 93US-00149503.
 PA (MERI) MERCK & CO INC.
 PI Elliott KJ, Ellis SB, Harpold NM;
 DR WPI; 2002-711528/77.
 DR N-FSDB; ABS54875.
 PT Identifying antagonists or agonists of human neuronal nicotinic
 PT acetylcholine receptors, by contacting recombinant cells with test
 PT compound, and measuring ion flux of cells or binding of compound to
 PT nNACHr.
 PS Claim 101; Col 59-64; 56pp; English.
 CC The invention relates to a method for identifying compounds that are
 CC antagonists or agonists of human neuronal nicotinic acetylcholine

CC receptors (nNACHrs), by contacting recombinant cells with a test compound
 CC and measuring ion flux, the electrophysiological response of the cells or
 CC binding of the test compound to the nNACHr. The recombinant cells are
 CC produced by transfection with a nucleic acid encoding at least one human
 CC nNACHr (alpha or beta) subunit, such that the cells express an nNACHr
 CC comprising one human subunit encoded by the transfected nucleic acid.
 CC This sequence represents the alpha 7 subunit of the human nNACHr
 CC polypeptide
 XX Sequence 502 AA;
 SQ Query Match 47.7%; Score 1258.5; DB 5; Length 502;
 Best Local Similarity 48.6%; Pred. No. 1.4e-122;
 Matches 250; Conservative 75; Mismatches 138; Indels 51; Gaps 7;
 QY 8 LALLA-LLPVSQGPHEKRLNALLANNTLRPVANESLEVRPLTQQIIVDEKN 66
 Db 10 LALAAALLHVSILQGEFQRLYKELVKNYNPLRPVANDSQPLTVYFSLQLQIMDVDEKN 69
 QY 67 QLLITNWLSEWLDYNLRWNSDEYGGVKDLRITPNKWLKPDVLMVNSADEGDFGTQTN 126
 Db 70 QVLTNINLQMSWTDHYLQNNVSEYGVKTVRPDQGIWKPDILLNYSADERDFATHTN 129
 QY 127 VVVRSGSCLYVPPGIFKSTCKMDIAWPFDDQHCMDKFGSWTYDGNQLDLVKDEAGD 186
 Db 130 VLVNSSGHCOYLPPIGIFKSCYDVRWFFDVQHCCKLFGSWSYGGSLDLQOE---AD 186
 QY 187 LSDFTNGEWYLGMPGKKNITTYACCPPEYVDVFTTIRRTLYFFNLIVPCVLISS 246
 Db 187 ISGYIPNGEWDLVGIPGKSERFYECKEYFPDVTFTVMRRRTLYYGLNLLIPCVLISA 246
 QY 247 MALLGFTLPDPSGEKLTGLVTLLSLTVFLNVAETLPQVSDAIPLGTYFNCIMFVNAS 306
 Db 247 LALLVFLPADSGEKISLIGITVLLSLTVFLLVAEIMPATSDSVPLIAQYFASTMTIIVGL 306
 QY 307 SVVLTVVVLYNHHRTADIHEMPQWIKSVFLOLPMILRMSRPGKTKITRKTIMMTRMREL 366
 Db 307 SVVTVIVLVYHHHDPDGGKMPKWTIRVILLNWCWFLMRKPGEDKVRPACQHKQRCSL 366
 QY 367 ELKERSKSLLANVLIDDDFRHGPPPNSTAGNL-----GPGC 407
 Db 367 ASVMSAVA-----PPP---ASNGNLLYIGRGLDGVHCVPTPDGSGVVC 407
 QY 408 SIFRTDFRSFVRPSTMEDV---GGGLGSHHRELHLILRELQFTARMKKADEAEELISDW 465
 Db 408 G-----RMACSPHDEHLHGQPPGDPDLAKILEEVRYIANFRCODESEAVCSEW 460
 QY 466 KFAAMVVDVDRFCLFVFTFTIATVAVLLSAPHII 499
 Db 461 KFAACVVDRLCLMAFSVFTIICIGILMSAPNFV 494
 RESULT 12
 ID ABG82435 standard; protein; 502 AA.
 AC ABG82435;
 DT 22-JAN-2003 (first entry)
 DE Human neuronal NACHr alpha7 subunit.
 KW Human; neuronal; nicotinic acetylcholine receptor; NACHr; drug screening;
 KW immunochemistry; NACHr alpha7 subunit; receptor.
 OS Homo sapiens.
 PN WO200259266-A2.
 PD 01-AUG-2002.
 PF 29-OCT-2001; 2001WO-US050985.
 XX

```
PR 01-NOV-2000; 2000US-00703951.
XX
XX (MERI ) MERCK & CO INC.
XX
XX Gillespie A, Claeps BO, Chavez-Noriega LE, Siegel R, Elliott KJ;
XX
XX WPI; 2002-698532/75.
DR N-PSDB; ABV73248.
XX
XX Cell comprising nucleic acids encoding human alpha and beta subunits of
PT neuronal nicotinic acetylcholine receptors, useful for in vitro screening
PT of a drug substance in a test system specific for humans.
XX
XX Example; Page 130-131; 143pp; English.
XX
XX The invention relates to a suitable host cell transfected with an
CC isolated nucleic acid molecule comprising a sequence of nucleotides or
CC ribonucleotides that encodes at least one alpha or beta subunit of a
CC human neuronal nicotinic acetylcholine receptor (nAChR). The compositions
CC and methods of the present invention, which provide a means to prepare
CC synthetic or recombinant receptors and receptor subunits that are
CC substantially free of contamination from many other receptor proteins,
CC are useful for observing the effect of a drug substance on a particular
CC subtype to perform in vitro screening of the drug substance in a test
CC system that is specific for humans. The antibodies can be used in
CC immunochemistry and for diagnostic and therapeutic applications. The
CC present sequence represents a human neuronal nAChR alpha7 subunit
XX
XX Sequence 502 AA;
SQ
Query Match 47.7%; Score 1258.5; DB 5; Length 502;
Best Local Similarity 48.6%; Pred. No. 1.4e-122;
Matches 250; Conservative 75; Mismatches 138; Indels 51; Gaps 7;
QY 8 LALLA-LLPVSEQGPHEKRLNALLANYNTLPRVANESEPLEVRFGLTQQIIVDDEKN 66
Db 10 LALAASLLHVSLOGEQFKLYKELVKYNPLERPVANDSQPLTVYFSLSLQIMDVDEKN 69
QY 67 QLLITNIWLSLENDYNLNRNDSEYGGVKDLRTIPNKLWKPDMVLYNSADEGDFGTQTN 126
Db 70 QVLTNTIWLQMSWDHYLQWNVSEYPGVKTRFPDQGIWKPDILLYNSADERFDPHTN 129
QY 127 VVVRSGSCLYVPPGIFKSTCKMDIAWFPDDQCHDKMGKSWTVDGNQLDLVLKDEAGD 186
Db 130 VLNSSGHQCLYLPPIGIFKSCYIDVRWFPDQVQCKLKFGSWSGWSLQMQE---AD 186
QY 187 LSDFITNGEYLIQMPGKKNITTYACCEPEYVDVTFITMIRRTLYYFFNLIVPCVLIS 246
Db 187 ISGYPNGEWDLVGIPGKRSERFYECCKEYPDPVTFVTMRRRTLYYGLNLLIPCVLISA 246
QY 247 MALLGFTLPDPSGEKLTGLVTILLSLTVFLNLVAETLPQVSDAIPLLGTFFNCIMFNWAS 306
Db 247 LALLVFLPADSGEKISLGTITVLLSLTVFVLLVAEIMPATSDSVPLIAQYFASTMTITVGL 306
QY 307 SVLTVVVLNVRHTADIHEMPOWIKSVFLQWLPWILMRSPGKIKTRKTIIMNTRMREL 366
Db 307 SVVTVIVLQVHHDDPGGKMPKTRVILLNWCWFLMRKRPBGDKVPACQHKQRCSL 366
QY 367 ELKRSKSLIANVLIDDDFRHGGPPPPNSTASTGNL-----GPGC 407
Db 367 ASVMSAVA-----PPP---ASGNLLYIGFRGLDGVCHCVPTDPSGVVC 407
QY 408 SIFTRFRSFRVSTMEDV---GGGLSHRHRELHLILRELOFITARMKKADEEELISDW 465
Db 408 G-----RWACSPTHDEHLHGQPPGDPDLAKILEEVRYIANRFRCDSEAVGSEW 460
QY 466 KFAAMVVDRECLFVFTLTFTIATVAVLLSAPHII 499
Db 461 KFAACVVDRLCLMAFSFTTICTIGILMSAPNFV 494
RESULT 13
ADA10874
```

ADA10874 standard; protein; 502 AA.

ADA10874;

06-NOV-2003 (first entry)

Human neuronal nicotinic acetylcholine receptor alpha 7 subunit #1.

alpha 7 subunit; human; neuronal nicotinic acetylcholine receptor; ligand-gated ion channel; synaptic transmission; gene therapy; transgenic; receptor.

Homo sapiens.

US6524789-B1.

25-FEB-2003.

07-JUN-1996; 96US-00660451.

07-JUN-1995; 95US-00484722.

(MERI) MERCK & CO INC.

Elliott KJ, Harpold MM;

WPI; 2003-511917/48.

N-PSDB; ADA10864.

New nucleic acid encoding an alpha-6 or a beta-3 subunit of a human neuronal nicotinic acetylcholine receptor (nAChR), useful for identifying compounds that modulate human neuronal nAChR activity.

Disclosure; Col 67-72; 63pp; English.

The invention relates to an isolated nucleic acid molecule comprising a nucleotide sequence encoding an alpha 6 or a beta 3 subunit of a human neuronal nicotinic acetylcholine receptor nAChR. nAChR's form ligand-gated ion channels that mediate synaptic transmissions between nerve and muscle and between neurons upon interaction with the neurotransmitter acetylcholine. The nucleic acid molecule is useful for identifying compounds that modulate human neuronal nAChR. The present sequence represents the amino acid sequence of the human neuronal nicotinic acetylcholine receptor, nAChR, alpha 7 subunit #1. Note: the present sequence is the sequence encoded by the nAChR alpha 7 subunit DNA ADA10864.

Sequence 502 AA;

Query Match 47.7%; Score 1258.5; DB 7; Length 502;

Best Local Similarity 48.6%; Pred. No. 1.4e-122;

Matches 250; Conservative 75; Mismatches 138; Indels 51; Gaps 7;

QY 8 LALLA-LLPVSEQGPHEKRLNALLANYNTLPRVANESEPLEVRFGLTQQIIVDDEKN 66

Db 10 LALAASLLHVSLOGEQFKLYKELVKYNPLERPVANDSQPLTVYFSLSLQIMDVDEKN 69

QY 67 QLLITNIWLSLENDYNLNRNDSEYGGVKDLRTIPNKLWKPDMVLYNSADEGDFGTQTN 126

Db 70 QVLTNTIWLQMSWDHYLQWNVSEYPGVKTRFPDQGIWKPDILLYNSADERFDPHTN 129

QY 127 VVVRSGSCLYVPPGIFKSTCKMDIAWFPDDQCHDKMGKSWTVDGNQLDLVLKDEAGD 186

Db 130 VLNSSGHQCLYLPPIGIFKSCYIDVRWFPDQVQCKLKFGSWSGWSLQMQE---AD 186

QY 187 LSDFITNGEYLIQMPGKKNITTYACCEPEYVDVTFITMIRRTLYYFFNLIVPCVLIS 246

Db 187 ISGYPNGEWDLVGIPGKRSERFYECCKEYPDPVTFVTMRRRTLYYGLNLLIPCVLISA 246

QY 247 MALLGFTLPDPSGEKLTGLVTILLSLTVFLNLVAETLPQVSDAIPLLGTFFNCIMFNWAS 306

Db 247 LALLVFLPADSGEKISLGTITVLLSLTVFVLLVAEIMPATSDSVPLIAQYFASTMTITVGL 306

QY 307 SVLTVVVLNVRHTADIHEMPOWIKSVFLQWLPWILMRSPGKIKTRKTIIMNTRMREL 366

Db 307 SVVTVIVLQVHHDDPGGKMPKTRVILLNWCWFLMRKRPBGDKVPACQHKQRCSL 366

QY 367 ELKRSKSLIANVLIDDDFRHGGPPPPNSTASTGNL-----GPGC 407

Db 367 ASVMSAVA-----PPP---ASGNLLYIGFRGLDGVCHCVPTDPSGVVC 407

QY 408 SIFTRFRSFRVSTMEDV---GGGLSHRHRELHLILRELOFITARMKKADEEELISDW 465

Db 408 G-----RWACSPTHDEHLHGQPPGDPDLAKILEEVRYIANRFRCDSEAVGSEW 460

QY 466 KFAAMVVDRECLFVFTLTFTIATVAVLLSAPHII 499

Db 461 KFAACVVDRLCLMAFSFTTICTIGILMSAPNFV 494

QY 307 SVLVTVVLYNHHRTADHMPQWIKSVFLQWLPWILMSRPGKKITRKTMMNTRMEL 366
Db 307 SVVTVIVLQYHHDDPGGKMPKTRVILLNKCWFLRMKPGEDKVPACQHKQRCSL 366
QY 367 ELKERSKSLANVLDDIDDDFRHGPPPPNSTASTGNL-----GPGC 407
Db 367 ASVEMSAVA-----PPP---ASGNLLYIGRGLDGVHCVPTDPSGVVC 407
QY 408 SIFRTDFRSFVRPSTMEDV--GGGLGSHHRELHLILRELQFITARMKKADEEALISDW 465
Db 408 G-----RMACSPTHDEHLHGQPPGDPDLAKILEEVRYIANFRFCQDESEAVCSEW 460
QY 466 KFAAMVVDVDFCLVFTLTIIATVAVLLSAPHII 499
Db 461 KFAACVVDRLCLMAFSVFTIICIGILMSAPNFV 494

RESULT 14

ADD47051
ID ADD47051 standard; protein; 502 AA.

AC ADD47051;

DT 29-JAN-2004 (first entry)

XX Human Protein NP_000737, SEQ ID NO 12739.

XX Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.

XX Homo sapiens.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346392P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GEO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX GENBANK; NP_000737.

XX New composition comprising two or more isolated polypeptides, useful for

XX preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating

CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 502 AA;

Query Match 47.7%; Score 1258.5; DB 7; Length 502;

Best Local Similarity 48.6%; Pred. No. 1.4e-122;

Matches 250; Conservative 75; Mismatches 138; Indels 51; Gaps 7;

QY 8 LALLA-LLPVSEQGPHEKRLNALLNANTLERPVANSEPLEVRFGLTLQIIVDDEKN 66

Db 10 LALAASLLHVSLOGEFQKLYKELVKNYNPLERPVANDSQPLTVVFSLSLQIMDVDEKN 69

QY 67 QLLITNIWLSLEWYNDYNLRWNSYGGVKDURITPNKWKPDVLMYNSADEGFGTYQTN 126

Db 70 QVLTNTIWLQMSWTDHYLQWNVSEYPGVKTVRFPDGGQIWKDPDILLNSADERFDATFTN 129

QY 127 VVVRSGGSLVVPPIFKSTCKMDIAWFPDDQHCMMKFGSWTYDGNQLDLVLKDEAGGD 186

Db 130 VLVNSGHCQYLPFGIFKSSCIYDVRWFPDQVCKLFGWSYGGWSLDIQMOE---AD 186

QY 187 LSDFITNGEWYLIQMPGKKNITTYACPEPYVDVTFTIMRRRTLYYFFNLIVPCVLIS 246

Db 187 ISGYIPNGEWDLVGIPGKRSERFYECKEPPYDVTFTVMTERRTLYYGLNLLIPCVLISA 246

QY 247 MALLGFTLPDPSGKLTGLVITLLSVFLNLVAETLPQVSDAIPLLGTYPNCIMFWVAS 306

Db 247 LALLVFLPADSGEKISLGITVLLSVFLNLVAETLPQVSDAIPLLGTYPNCIMFWVAS 306

QY 307 SVLTVVLYNHHRTADHMPQWIKSVFLQWLPWILMSRPGKKITRKTMMNTRMEL 366

Db 307 SVVTVIVLQYHHDDPGGKMPKTRVILLNKCWFLRMKPGEDKVPACQHKQRCSL 366

QY 367 ELKERSKSLANVLDDIDDDFRHGPPPPNSTASTGNL-----GPGC 407

Db 367 ASVEMSAVA-----PPP---ASGNLLYIGRGLDGVHCVPTDPSGVVC 407

QY 408 SIFRTDFRSFVRPSTMEDV--GGGLGSHHRELHLILRELQFITARMKKADEEALISDW 465

Db 408 G-----RMACSPTHDEHLHGQPPGDPDLAKILEEVRYIANFRFCQDESEAVCSEW 460

QY 466 KFAAMVVDVDFCLVFTLTIIATVAVLLSAPHII 499

Db 461 KFAACVVDRLCLMAFSVFTIICIGILMSAPNFV 494

RESULT 15

ADE57310

ID ADE57310 standard; protein; 502 AA.

AC ADE57310;

DT 29-JAN-2004 (first entry)

XX Human Protein P36544, SEQ ID NO 3171.

XX Human; pain; neuronal tissue; gene therapy;

XX spinal segmental nerve injury; chronic constriction injury; CCI;

XX spared nerve injury; SNI; Chung.

XX Homo sapiens.

XX WO2003016475-A2.

XX 27-FEB-2003.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 8, 2004, 06:45:49 ; Search time 5514.12 Seconds
(without alignments)
3938.047 Million cell updates/sec

Title: US-09-303-232-6

Perfect score: 2640

Sequence: 1 MAPMAALALLLPVSEQ.....LFTIATVALLSAPHIVQ 501

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh

-Q=/cgn2_1/USFto.spool/US09303232/runat_07052004_101110_23883/app query.fasta_1.2261

-DB=genembl -Qfmt=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0

-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45

-DOALIGN=200 -THR_SCORE=pc -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL

-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US09303232@cgn_1_1_8225@runat_07052004_101110_23883 -NCPU=6 -ICPU=3

-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_man.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2640	100.0	3029	3	AF143847 Heliothis
2	2640	100.0	3109	6	E58348 Nucleic aci
3	2640	100.0	3109	6	AX009614 Sequence
4	2194	83.1	2023	3	AF321447 Drosophil
5	2181	82.6	2023	3	AF321446 Drosophil
6	2177	82.5	1665	3	DME554209
7	2177	82.5	1699	3	BT011147
8	2177	82.5	2023	3	AF321445
9	2176.5	82.4	2068	3	AF321448
10	2156.5	81.7	2110	3	AF321449
11	1803.5	68.3	3629	3	AF143846
12	1803.5	68.3	3700	6	AX009612
13	1803.5	68.3	3701	6	E58347
14	1801.5	68.2	1683	3	DME554210
15	1800.5	68.2	2907	3	AF272778
16	1671	63.3	2834	3	AY036613
17	1609	60.9	2886	6	E58346
18	1609	60.9	2886	6	AX009610
19	1260	47.7	2087	9	HSARA7A
20	1258.5	47.7	1509	6	AX054567
21	1258.5	47.7	1509	9	HSNACHRA7
22	1258.5	47.7	1876	6	AR055255
23	1258.5	47.7	1876	6	AR071403
24	1258.5	47.7	1876	6	AR173187
25	1258.5	47.7	1876	6	AR224030
26	1258.5	47.7	1876	6	AR282833
27	1258.5	47.7	1876	6	AX719088
28	1258.5	47.7	1876	9	HSU62436
29	1258.5	47.7	2088	10	AF225980
30	1255.5	47.6	1848	10	MUSNARS
31	1255.5	47.6	2106	10	RATNARAD
32	1254.5	47.5	1590	6	AR261850
33	1254.5	47.5	2107	6	BD023656
34	1254.5	47.5	2107	9	AF486623
35	1253	47.5	2037	5	GGNNARA7A
36	1253	47.5	2696	5	GGA7NAREC
37	1252.5	47.4	1509	6	AX054575
38	1252.5	47.4	1916	9	AF385585
39	1252.5	47.4	1977	9	HSU40583
40	1250	47.3	1559	9	HUMA7NAR
41	1248.5	47.3	1509	6	AX054577
42	1246.5	47.2	3030	10	S53987
43	1245	47.2	1668	5	AY298752
44	1242.5	47.1	1509	6	AX054579
45	1240.5	47.0	1551	4	BTA7

ALIGNMENTS

RESULT 1

AF143847 3029 bp mRNA linear INV 27-MAY-1999
 LOCUS Heliothis virescens putative nicotinic acetylcholine receptor alpha
 DEFINITION 7-2 subunit mRNA, complete cds.
 ACCESSION AF143847
 VERSION AF143847.1 GI:4895006
 KEYWORDS Heliothis virescens (tobacco budworm)
 SOURCE Heliothis virescens
 ORGANISM Rukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea; Noctuidae; Heliothinae; Heliothis.
 REFERENCE 1 (bases 1 to 3029)
 AUTHORS Schulte, T., Oellers, N. and Adamczewski, M.
 TITLE Putative alpha subunits of insect nicotinic acetylcholine receptors more similar to vertebrate alpha 7 subunits and C. elegans Ce21 than to other insect nicotinic acetylcholine receptor alpha subunits
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 3029)
 AUTHORS Schulte, T., Oellers, N. and Adamczewski, M.
 TITLE Direct Submission
 JOURNAL Submitted (19-APR-1999) ZF-BTB, Bayer AG, Bldg. Q 18., Leverkusen 51368, Germany
 FEATURES
 source Location/Qualifiers
 1..3029
 /organism="Heliothis virescens"
 /mol_type="mRNA"
 /db_xref="taxon:7102"
 95..1600
 /note="hvnachra7-2"
 /codon_start=1
 /product="putative nicotinic acetylcholine receptor alpha 7-2 subunit"
 /protein_id="AAD32698.1"
 /db_xref="GI:4895007"
 /translation="MAPMLAALALLIPVSEQPHEKRLINALLANVTILRVPVANE
 SEPLEVFLGLTQOI IDVDENQQLITNIWLEWNDVNLWNSDEYGVKDLRIPTN
 KLWPDVILNSADEGFDGTQTNVVRSGSCLYVPGIFKSTCKMDIAWFPDDOH
 CDMKFGWTYDGNQDILVKDEAGDLSDFITINGEWYLI GMPGKNITITYACPEPYV
 DVFTIMIRRTLYFNLIVPCLVISMALLGFTLPDPSGEKLTGLVITLLSUTVEL
 NLVATLPOVSDATPLKGTFCNFMFVASSVLTVVYVNLNYHRTADIHMPQWIKSV
 FLOWLILKMSRPGKIKTRKIMNTRMRLEKRSKSLANLVIDDDFRHGP
 PPNSTAGTGNLGPSCSIFRTDFRSFVRPSTMEDVGGGLGSHHRELHLIRELQFIITA
 RMKKADEAELISDWKFAAMVDRFCULFVFTLITIAVAVLLSAPHLIQV"
 ORIGIN
 Alignment Scores:
 Pred. No.: 7,52e-245 Length: 3029
 Score: 2640.00 Matches: 501
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 3 Gaps: 0
 US-09-303-232-6 (1-501) x AF143847 (1-3029)
 QY 1 MetAlaProMetLeuAlaAlaLeuAlaLeuAlaLeuProValSerGluGlnGly 20
 Db 95 ATGGCCCCATATGTTGGCGCCCTTGGCGCTGTGGCTTTGGTCCCGTATCGGAGCAAGGT 154
 QY 21 ProHisGluLysArgLeuLeuAsnAlaLeuLeuAlaAsnTyrAsnThrLeuGluArgPro 40
 Db 155 CCTACGAGAGAGAGACTCTCGACCGCGTGTGGCGAACTACACACCCTCGGAGCAGCG 214
 QY 41 ValAlaAsnGluSerGluProLeuGluValArgPheGlyLeuThrLeuGlnGlnIleLe 60
 Db 215 GTGGCCACACGAGCGAACCGGTAGAGGTCAGGTCAGGTCGGCTTGCACCTTCAGCAATCAT 274
 QY 61 AspValAspGluLysAsnGlnLeuLeuIleThrAsnIleThrLeuSerLeuGluTrpAsn 80
 Db 275 GACGTGGAGGAGAAATCAACTACTTATAACCAATATATGGCTGCTGCTGGAGTGAAT 334

QY 81 AspTyrAsnLeuArgTrpAsnAspSerGluTyrGlyValLysAspLeuArgIleThr 100
 Db 335 GACTACACCTGAGTGGACGACGAGTATGGCGGGTCAAGGACCTCAGGATCAGC 394
 QY 101 ProAsnLysLeuTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAsp 120
 Db 395 CCCAACCAAGTTGTGGAAGCGGACGCTCTTATGTATATAGTGTCTGACGAGGGTTTTC 454
 QY 121 GlyThrTyrGlnThrAsnValValArgSerGlyValSerCysLeuTyrValProPro 140
 Db 455 GGGACCTTACCAGACCAACGCTGTGTGAGAGCGCGGCGAGTCTCTGTAGTGCACCT 514
 QY 141 GlyIlePheLysSerThrCysLysMetAspIleAlaTrpPheProPheAspAspGlnHis 160
 Db 515 GGCATATTCAGAGCACATGCAAGATGACATCGCGTGGTTCCTTCGACGACCAACAC 574
 QY 161 CysAspMetLysPheGlySerTrpThrTyrAspGlyAsnGlnLeuAspLeuValLeuLys 180
 Db 575 TGTGATATGAAGTTCCGTAGCTGGACATATGACGCGCAATCAGTTGGATCTGGTCTAAA 634
 QY 181 AspGluAlaGlyGlyAspLeuSerAspPheIleThrAsnGlyGluTyrTyrLeuIleGly 200
 Db 635 GATCAGCGACGCGCGGATCTATCGGACTTCATAACAATGGGAGTGGTATCTAATAGGA 694
 QY 201 MetProGlyLysLysAsnThrIleThrTyrAlaCysCysProGluProTyrValAspVal 220
 Db 695 ATGCCAGGCAAAAGAACACAAATAACATACGCGTGTCTGCCCGAGCCCTACGTGACGTC 754
 QY 221 ThrPheThrIleMetIleArgArgArgThrLeuTyrTyrPhePheAsnLeuIleValPro 240
 Db 755 ACCTTCCACCATCATGATAAGAGACGAACCTTGTACTACTTCTTCAACCTGATCGTCCG 814
 QY 241 CysValLeuIleSerSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGlu 260
 Db 815 TGGCGTGTGATCTCATCGATGGCACTCTCGGCTTCACATGCCACACGACTCCGGAGAG 874
 QY 261 LysLeuThrLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnLeuValAla 280
 Db 875 AAACCTCACACTTGCAGTCACATATTCTTCTATCGTGAGCGGTGTCTCTCAACCTGTGATCC 934
 QY 281 GluThrLeuProGlnValSerAspAlaIleProLeuLeuGlyThrTyrPheAsnCysIle 300
 Db 935 GAGACCTCGCCACAGGTCCTCGCGCTATCCCTCTGTAGGAGCGTACTTCAATTGCATC 994
 QY 301 MetPheMetValAlaSerSerValValLeuThrValValLeuAsnTyrHisArg 320
 Db 995 ATGTTTCATGGTAGCGTCTGTGTGTACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1054
 QY 321 ThrAlaAspIleHisGluMetProGlnTrpIleLysSerValPheLeuGlnTrpLeuPro 340
 Db 1055 ACAGCTGATATACATGAAATGCCACAGTGGATAAAATCAGTATTCTTCAATGGTGGCA 1114
 QY 341 TrpIleLeuArgMetSerArgProGlyLysIleThrArgLysThrIleMetMetAsn 360
 Db 1115 TGGATACTCGAATGTGAGGCCAGGGAAGAGATCACCGAAGAACTATTAATGATGAAC 1174
 QY 361 ThrArgMetArgGluLeuGluLeuLysGluArgSerSerLysSerLeuLeuAlaAsnVal 380
 Db 1175 ACGAGGATGAGGAGCTGGAACCTGAAGGAGAGTCTCGAAGTCTCTGCTGGCGAATGTT 1234
 QY 381 LeuAspIleAspAspPheArgHisGlyProProProProAsnSerThrAlaSerThr 400
 Db 1235 CTAGATATTGATGATGACTTCAGACACGCGCTCTCGCTCTCTAAACAGTACTGCCTCGAC 1294
 QY 401 GlyAsnLeuGlyProGlyCysSerIlePheArgThrAspPheArgArgSerPheValArg 420
 Db 1295 GGGAAATTTGGACCTGGGTGTCAATATTCGCGACGGAATTTCCGTCGTCGTCTGTCGTGT 1354
 QY 421 ProSerThrMetGluAspValGlyGlyLeuGlySerHisHisArgGluLeuHisLeu 440
 Db 1355 CCGTCCAGCATGGAGACCTGGCGCGCGGTGGGTGACCCACCATCGGAGCTGCACCTC 1414
 QY 441 IleLeuArgGluLeuGlnPheIleThrAlaArgMetLysLysAlaAspGluAlaGlu 460

Db 1415 ATACTGAGAGGCTGCACTTTCATACCGCCGAGTGAAGAGCGCTGATGAGGAGCCGAG 1474
 QY 461 LeuileSerAspTrpLysPheAlaAlaMetValValAspArgPheCysLeuPheValPhe 480
 Db 1475 CTGATCAGCGACTGGAAGTTTGCTGCGATGCTTGTGATAGGTTTTCCTGCTGCTTC 1534
 QY 481 ThrLeuPheThrIleAlaThrValAlaValLeuLeuSerAlaProHisIleIleVal 500
 Db 1535 ACATTTTCAATCATCGGACAGTAGCTGCTCTTATCGGACCGCATATCATCGTG 1594
 QY 501 Gln 501
 Db 1595 CAA 1597

RESULT 2

LOCUS E58348 3109 bp DNA linear PAT 18-JUN-2001
 DEFINITION Nucleic acid encoding insect acetyl choline receptor subunit.
 ACCESSION E58348
 VERSION E58348.1 GI:13019347
 KEYWORDS JP 2000023680-A/3.
 SOURCE Heliothis virescens (tobacco budworm)
 ORGANISM Heliothis virescens
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 Noctuidae; Noctuidae; Heliothinae; Heliothis.
 1 (bases 1 to 3109)
 Martin, A., Nadja, E. and Thomas, S.
 Nucleic acid encoding insect acetyl choline receptor subunit
 Patent: JP 2000023680-A 3 25-JAN-2000;
 BAYER AG

COMMENT

OS Heliothis virescens
 PN JP 2000023680-A/3
 PD 25-JAN-2000
 PF 26-APR-1999 JP 1999118159
 PR 04-MAY-1998 DE 19819829.9
 PI MARTIN ADAMUTSUEUSUKI, NADJA, ERASU, THOMAS SCHULTE PC
 C12N15/09, A01K67/033, C07K14/705, C07K16/28, C12N1/21, C12N5/10, PC
 C12Q1/68,
 PC G01N33/15, G01N33/50// (C12N1/21, C12R1/19), C12N15/00, C12N5/00 CC

FEATURES
source

FH Key Location/Qualifiers
 FT CDS Location/Qualifiers
 1..3109
 /organism="Heliothis virescens"
 /mol_type="genomic DNA"
 /db_xref="taxon:7102"

ORIGIN

Alignment Scores:
 Pred. No.: 7.79e-245 Length: 3109
 Score: 2640.00 Matches: 501
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-303-232-6 (1-501) x E58348 (1-3109)

QY 1 MetAlaProMetLeuAlaAlaLeuAlaLeuAlaLeuProValSerGluGlnGly 20
 Db 95 ATGGCCCTATATGTGGCGGCTTGGCGCTGCTGGCTTTGCTGCCCTATCGGAGCAAGGT 154
 QY 21 ProHisGluLysArgLeuLeuAsnAlaLeuLeuAlaAsnTyrAsnThrLeuGluArgPro 40
 Db 155 CCTCAGAGAGAGACTCTCTGAACGCTTCTGCGCACTACACACCCTGGAGCGACCG 214
 QY 41 ValAlaAsnGluSerGluProLeuGluValArgPheGlyLeuThrLeuGlnGlnIle 60
 Db 215 GTGGCCAACGAGACGACCGCTAGAGGTGAGGTTCGGCTTGACCTTGACCAAAATCAIT 274

QY 61 AspValAspGluLysAsnGlnLeuLeuIleThrAsnIleTrpLeuSerLeuGluTrpAsn 80
 Db 275 GAGCTGGACGAGAGAAATCAACTACTTATAACCAATATATGGCTGTGCTGGAGTGAAT 334
 QY 81 AspTyrAsnLeuArgTrpAsnAspSerGluTyrGlyGlyValLysAspLeuArgIleThr 100
 Db 335 GACTACACCTGAGGTGGAAACGACGAGAGTATGGCGGGGTCAAGGACCTCAGGATCACG 394
 QY 101 ProAsnLysLeuTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAsp 120
 Db 395 CCCAACAAAGTTGTGGAAGCCGACGCTCTTATGTATAATAGTGTCTGACGAGGGTTTGAC 454
 QY 121 GlyThrTyrGlnThrAsnValValValArgSerGlyGlySerCysLeuTyrValProPro 140
 Db 455 GGGACCTTACCAGACCAACGCTGGTGGTCAGAGCGCGGCGAGTTGCCCTGTACGTGCCACCT 514
 QY 141 GlyIlePheLysSerThrCysLysMetAspIleAlaTrpPheProPheAspAspGlnHis 160
 Db 515 GGCATATTCAAGAGCACATGCAAGATGGACATCGCTGGTTCCTTCGACGACCAACAC 574
 QY 161 CysAspMetLysPheGlySerTrpThrTyrAspGlyAsnGlnLeuAspLeuValLys 180
 Db 575 TGTGATATGAAGTTCCGTAGCTGCACATATGACGGCAATCAGTTGGATCTGGTCTAAA 634
 QY 181 AspGluAlaGlyGlyAspLeuSerAspPheIleThrAsnGlyGluTrpTyrLeuIleGly 200
 Db 635 GATGAGCGAGCGCGGCGATCTATCGGACTTCATACAAATGGGAGTGTATCTAATAGGA 694
 QY 201 MetProGlyLysLysAsnThrIleThrTyrAlaCysCysProGluProTyrValAspVal 220
 Db 695 ATGCCAGGCAAAAAGAACACATACATAGCGGTGTGCCCGGAGCCCTACGTGGAGCGTC 754
 QY 221 ThrPheThrIleMetIleArgArgThrLeuTyrTyrPhePheAsnLeuIleValPro 240
 Db 755 ACCTTCACCATCATGATAGAGACGAAACCTTCTACTACTTCTCAACCTGATCGTCCCG 814
 QY 241 CysValIleSerSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGlu 260
 Db 815 TGGTGTCTGATCTCATCGATGGCACTCCCTGGCTTCACACTGCCACGAGCTCCGAGAG 874
 QY 261 LysLeuThrLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnLeuValAla 280
 Db 875 AAACACACACTGGAGTCATCTTCTCTATCGCTGCGGTGTCTCAACCTGGTAGCC 934
 QY 281 GluThrLeuProGlnValSerAspAlaIleProLeuLeuGlyThrTyrPheAsnCysIle 300
 Db 935 GAGACCTCGCCACAGGTCTCCGACGCTATCCCCCTGTAGGGAGCTACTTCAATTGCATC 994
 QY 301 MetPheMetValAlaSerSerValValLeuThrValValLeuAsnTyrHisHisArg 320
 Db 995 ATGTTTCATGGTAGCGTCGCTGTGGTACTGACTGTGGTGTACTCAATTACCACCATCGA 1054
 QY 321 ThrAlaAspIleHisGluMetProGlnTrpIleLysSerValPheLeuGlnTrpLeuPro 340
 Db 1055 ACAGCTGATATACATGAATGCCACAGTGGATAAATCAGTATTCCTACNATGGTGCCA 1114
 QY 341 TrpIleLeuArgMetSerArgProGlyLysLysIleThrArgLysThrIleMetMetAsn 360
 Db 1115 TGGATACTCGAATGTCCAGGCCAGGGAAGAGATCACCAGGAAGACTATAATGATGAAC 1174
 QY 361 ThrArgMetArgGluLeuLeuLysGluArgSerSerLysSerLeuLeuAlaAsnVal 380
 Db 1175 ACAGGATGAGGAGGTGGAACCTGAAGGAGAGGTGCTCGAAGTCTTGTGCGCAATGTT 1234
 QY 381 LeuAspIleAspAspPheArgHisGlyProProProProAsnSerThrAlaSerThr 400
 Db 1235 CTAGATATTGATGATGATCTTCAGACACGCGCTCCGCTCTCTAACAGTACTGCCTCGACC 1294
 QY 401 GlyAsnLeuGlyProGlyCysSerIlePheArgThrAspPheArgArgSerPheValArg 420
 Db 1295 GGGAAATTTGGGACCTGGGTGCTCAATATTCGACGAGATTTCCGTGGTGGTCCGTCCGT 1354
 QY 421 ProSerThrMetGluAspValGlyGlyGlyLeuGlySerHisHisArgGluLeuHisLeu 440

```
Db      1355 CCGTCCACGATGGAGACGTGGGCGGCGGTGGTAGCCACCACCTCGGAGTCACCTC 1414
Qy      441 ILeuArgGluLeuGlnPheIleThrAlaArgMetLysLysAlaAspGluGluAlaGlu 460
Db      1415 ATACTGAGAGCTGCAGTTTCATCAGCGCAGGATGAAGAGGCTGATGAGGAAGCCGAG 1474
Qy      461 LeuIleSerAspTrpLysPheAlaAlaMetValValAspArgPheCysLeuPheValPhe 480
Db      1475 CTGATCAGCACTGGAAGTTTGTGCGATGGTTGTTGATAGTTTTCCTGTTTCGTGTC 1534
Qy      481 ThrLeuPheThrIleAlaThrValAlaValLeuLeuSerAlaProHisIleIleVal 500
Db      1535 ACATTTTCACAAATCATCGGACAGTAGTCTCTGTTATCGGACCGCATATCATCGT 1594
Qy      501 Gln 501
Db      1595 CAA 1597

RESULT 3
AX009614
LOCUS      AX009614          3109 bp      DNA      linear      PAT 06-SEP-2000
DEFINITION Sequence 5 from Patent EP0962528.
ACCESSION  AX009614
VERSION     AX009614.1  GI:9996846
KEYWORDS
SOURCE      Heliothis virescens (tobacco budworm)
ORGANISM    Heliothis virescens
            Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
            Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
            Noctuoidea; Noctuidae; Heliothinae; Heliothis.
REFERENCE   1
            Adamczewski,M.D., Schulte,T.D. and Oellers,N.D.
            Nucleic acids encoding acetylcholin-receptor subunits from insects
            Patent: EP 0962528-A 5 08-DEC-1999;
            BAYER AG (DE)
FEATURES
            source
            1. 3109
                /organism="Heliothis virescens"
                /mol_type="unassigned DNA"
                /db_xref="taxon:7102"
                95..1600
                /note="unnamed protein product"
                /codon_start=1
                /protein_id="CAC07501.1"
                /db_xref="GI:9996847"
                /db_xref="FEMTREMEL:CAC07501"
                /translation="MAPMLAALALLLPVSEQQPHEKRLNALLANNYTLERPVANE
                SEPLEVRFGLTQGIIDVDEKNQLLIINIWLSEWNYNLRWNSDSEYGGVKDLRITPN
                KLMKPDVLMYNSADGFDGTQTVNVRSGGSLYVPPGIFKSTCKMDIAWPFDDOH
                CDMKFGSWTDGNOLDLVLKDEAGDLSDRITNGEWYLGMPGKKNTITVACCPPEYV
                DVTFTIMRRRTLYFFNLIVPCVLISMAILGFTLPDSEKLTGLVILLSLTVEL
                NLVAETLPQVSDALPLLGTYFNCIMFWVASSVLTIVVLYNTHRTADIHENPQWKS
                FLOWLPILMRSGKKITRTKIMTMRELEKRSKSLANLVIDDDDFRHPGP
                PPNSTAGNLPGCCSIFRTDFRSFVRPTMEDVGGGLGSHRELHLILRELOQFITA
                RMKKADEAELISDWKFAAMVVDVDFCLFVFTLFTIATVAVLLSAPHILVQ"
ORIGIN
Alignment Scores:
Pred. No.:      7,79e-245      Length:      3109
Score:          2640.00      Matches:      501
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:     100.00%      Indels:      0
DB:              6      Gaps:      0

US-09-303-232-6 (1-501) x AX009614 (1-3109)
Qy      1 MetAlaProMetLeuAlaAlaLeuAlaLeuLeuAlaLeuProValSerGluGlnGly 20
Db      95 ATGGCCCTTAAGTTGGCGGCTTGGCGTGTCTGGCTTGTGGCCGATCGGAGCAAGGT 154
Qy      21 ProHisGluLysArgLeuLeuAsnAlaLeuLeuAlaAsnTyrAsnThrLeuGluArgPro 40
```

```
Db      155 CCTCAGAGAGAGACTCTCTGAACGCGTGTGTCGGAACATACAACACCTTGAGCGACCG 214
Qy      41 ValAlaAsnGluSerGluProLeuGluValArgPheGlyLeuThrLeuGlnGlnIleIle 60
Db      215 GTGGCCAAACGAGAGCGAAGCGCTAGAGGTTCAGGTTCGGCTTGACCTTGCAGCAATCATT 274
Qy      61 AspValAspGluLysAsnGlnLeuLeuIleThrAsnIleTrpLeuSerLeuGluTrpAsn 80
Db      275 GACGTGAGCAGAGAAGAAATCAACTACTTATTAACCAATATATGGCTGTCTGTGGAGTGAAT 334
Qy      81 AspTyrAsnLeuArgTrpAsnAspSerGluTyrGlyGlyValLysAspLeuArgIleThr 100
Db      335 GACTTACAACCTGAGTGGAAACGACGAGATGATGGCGGGTCAAGGACCTCAGGATCAG 394
Qy      101 ProAsnLysLeuTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAsp 120
Db      395 CCCAACAAAGTTGTGGAAGCCGACGCTCTTATGTATATATAGTGTGACAGGGGTTTGAC 454
Qy      121 GlyThrTyrGlnThrAsnValValArgSerGlyGlySerCysLeuTyrValProPro 140
Db      455 GGGACCTTACAGACCAACGTTGGTGCAGAAAGCGCGGACGATTGCCCTGTAGTGCACCT 514
Qy      141 GlyIlePheLysSerThrCysLysMetAspIleAlaTrpPhePhePropheAspGlnHis 160
Db      515 GGCATATTCAAGAGCACATGCAAGATGCATCGCTGGTGTTCCTTCGACGACCAACAC 574
Qy      161 CysAspMetLysPheGlySerTrpThrTyrAspGlyAsnGlnLeuAspLeuValLeuLys 180
Db      575 TGTGATATGAAGTTTCGGTAGCTGACATATGACGCAATCAGTTGGATCTGGTCTAAAA 634
Qy      181 AspGluAlaGlyLysAspLeuSerAspPheIleThrAsnGlyGluTyrTyrLeuIleGly 200
Db      635 GATGAGCGAGCGCGCGATCTATCGCATTTATACAAATGGGAGTGGTATCTAATAGGA 694
Qy      201 MetProGlyLysLysAsnThrIleThrTyrAlaCysCysProGluProTyrValAspVal 220
Db      695 ATGCCAGGCAAAAAGAAACAAATAACATACGCGTCTGCCCGAGCCCTACGTGACGCTC 754
Qy      221 ThrPheThrIleMetIleArgArgArgThrLeuTyrTyrPhePheAsnLeuValPro 240
Db      755 ACCTTCACCATCATGATGGCACTCTCGGCTTCACACTGCCACGACGATCCCGGAGAG 814
Qy      241 CysValLeuIleSerSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGlu 260
Db      815 TGCCTGTGATCTCATCGATGGCACTCTCGGCTTCACACTGCCACGACGATCCCGGAGAG 874
Qy      261 LysLeuThrLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnLeuValAla 280
Db      875 AAACCTCACACTTGGAGTCACTATTCTTCTATCGCTGACGGTGTTCCTCAACCTGTAGCC 934
Qy      281 GluThrLeuProGlnValSerAspAlaIleProLeuLeuGlyThrTyrPheAsnCysIle 300
Db      935 GAGACCTGCCACAGGCTCCGACGCTATCCCTCTGTAGGAGCGTACTTCAATGTCATC 994
Qy      301 MetPheMetValAlaSerSerValValLeuThrValValLeuAsnTyrHisIleSar 320
Db      995 ATGTTTCATGTTAGCTGCTCTGTGTAAGTCTGCTGCTGCTCAATTACCATCGA 1054
Qy      321 ThrAlaAspIleHisGluMetProGlnTrpIleLysSerValPheLeuGlnTrpLeuPro 340
Db      1055 ACAGCTGATATACATGAATAATGCCACAGTGGATAAAATCAGTATTCTTACAAATGGTTGCA 1114
Qy      341 TrpIleLeuArgMetSerArgProGlyLysLysIleThrArgLysThrIleMetMetAsn 360
Db      1115 TGGATACGCAATGTGAGGCCAGGGAAGATCACCAGGAGAGACTATATATGATGAC 1174
Qy      361 ThrArgMetArgGluLeuGluLeuLysGluArgSerSerLysSerLeuLeuAlaAsnVal 380
Db      1175 ACGAGGATGAGGAGCTGGAATCAAGGAGAGGTCGTGAAAGTCCTTCTGCGCAATGTT 1234
Qy      381 LeuAspIleAspAspPheArgHisGlyProProProProAsnSerThrAlaSerThr 400
```


Qy 196 TrpTyrLeuIleGlyMetProGlyLysLysAsnThrIleThrTyrAlaCysCysProGlu 215
Db 985 TGGTACTTGCTGTCATCGCGGAAAGAAATACAGTAGTCTACGCTGCTGCCAGAA 1044
Qy 216 ProTyrValAspValThrIleMetIleArgArgThrLeuTyrTyrPhePhe 235
Db 1045 CCATATGCGATATCACCTTTACTATACAAATTCGTCGCGTACATTATATTTTTC 1104
Qy 236 AsnLeuIleValProCysValLeuIleSerSerMetAlaLeuLeuGlyPheThrLeuPro 255
Db 1105 AATTATATCGCGCATGTGCTAATCTCATCGATGCGCTTACGCGCTTACATTCGCG 1164
Qy 256 ProAspSerGlyGlyLysLeuThrLeuGlyValThrIleLeuLeuSerLeuThrValPhe 275
Db 1165 CCGGATTCGCGGAGAACTGACGCTGGAGTTACAATCTTCTATCGCTACAGTGTTC 1224
Qy 276 LeuAsnLeuValAlaGluThrLeuProGlnValSerAspAlaIleProLeuLeuGlyThr 295
Db 1225 CTCACCTTGTAGCTGAGACATTGCCCAAGTATCTGATGCAATCCCTTGTAGGCACC 1284
Qy 296 TyrPheAsnCysIleMetPheMetValAlaSerSerValLeuThrValValValLeu 315
Db 1285 TACTTCAATTGTCATCATGTCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1344
Qy 316 AsnTyrHisArgThrAlaAspIleHisGluMetProGlnTrpIleLysSerValPhe 335
Db 1345 AACTACCAACATCGCACAGCGGACATTCACAGATGCCCGTGGATCAAGTCGTTTC 1404
Qy 336 LeuGlnTrpLeuProTrpIleLeuArgMetSerArgProGlyLysLysIleThrArgLys 355
Db 1405 CTACAATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1464
Qy 356 ThrIleMetMetAsnThrArgMetArgGluLeuLeuLysGluArgSerSerLysSer 375
Db 1465 ACAATACTATTAAAGCAATCGCATGAAGAGCTGAGCTAAAGAGCGCTCTCCAAATCC 1524
Qy 376 LeuLeuAlaAsnValLeuAspIleAspAspPheArgHisGlyProProProProAsn 395
Db 1525 CTGCTGGCAATGCTCTCGACATCGACGACACTTCGCGAC----- 1566
Qy 396 SerThrAlaSerThrGlyAsnLeuGlyProGlyCysSerIlePheArgThrAspPheArg 415
Db 1567 ---ACAATATCTGGCTCCAAACCGCCATTCGCTCGTCTG----- 1602
Qy 416 ArgSerPheValArgProSerThrMetGluAspValGlyGlyLeuGlySerHisHis 435
Db 1603 GCCAGCTTCGCTCGGCCACACGCTGGAGGAGCATCACGCGCATCGCTGCAATCAC 1662
Qy 436 ArgGluLeuHisLeuIleLeuArgGluLeuGlnPheIleThrAlaArgMetLysLysAla 455
Db 1663 AAAGATCTTCACTTAATCTCAAGAAATTCGAATTTATTACGCGCGGATCGCAAGCT 1722
Qy 456 AspGluGluAlaGluLeuSerAspTrpLysPheAlaAlaMetValValAspArgPhe 475
Db 1723 GACGACGAAGCGGAATTCATCGCGATTCGGAATTCGCGCAATGTTGTGATAGATT 1782
Qy 476 CysLeuPheValPheThrLeuPheThrIleIleAlaThrValAlaValLeuLeuSerAla 495
Db 1783 TGTTTAATGTTTTCACGCTCTTACGAATTAATGCAACGGTTACGCTGCTCTCCGCT 1842
Qy 496 ProHisIleIleValGln 501
Db 1843 CCGCACATAATCGTGCAA 1860
RESULT 5
AF321446 2023 bp mRNA linear INV 29-APR-2002
LOCUS Drosophila melanogaster nicotinic acetylcholine receptor Dalpha6
DEFINITION subunit variant type II (nAcRalpha-30D) mRNA, complete cds,
alternatively spliced.
ACCESSION AF321446
VERSION AF321446.1 GI:20152846

KEYWORDS

SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 2023)
AUTHORS Grauso, M., Reenan, R.A., Culetto, E. and Sattelle, D.B.
TITLE Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
Dalpha5, Dalpha6 and Dalpha7, in Drosophila melanogaster Identify a
New and Highly Conserved Target of Adenosine Deaminase Acting on
RNA-Mediated A-to-I Pre-mRNA Editing
Genetics 160 (4), 1519-1533 (2002)
JOURNAL 21969411
MEDLINE 11973307
PUBMED 2 (bases 1 to 2023)
REFERENCE 11973307
AUTHORS Grauso, M. and Sattelle, D.B.
TITLE Direct Submission
JOURNAL Submitted (15-NOV-2000) Human Anatomy and Genetics-Functional
Genetics Unit, University of Oxford, South Park Road, Oxford OX1
3QX, UK
FEATURES
source Location/Qualifiers
1. .2023
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/chromosome="2"
/map="30D1"
/dev_stage="embryo"
1. .2023
/gene="nAcRalpha-30D"
379. .1863
/gene="nAcRalpha-30D"
/note="ion channel; neurotransmitter transmembrane
receptor; alternatively spliced; contains exons 3b and 8a"
/codon_start=1
/product="nicotinic acetylcholine receptor Dalpha6 subunit
variant type II"
/protein_id="AA013393.1"
/db_xref="GI:20152847"
/translation="MDSPLASLSLFLVLLIFLAIKESCOCGHEKLLNHLSTVNTL
ERPVAESELEVKFGTLTQOIIDVDKRNQLITNMLSLWNDYNLRWNETEGGVK
DLRITPKMKPDVIMYNSADEGPDGYHTNIVVKGSGCLYVPPGIFKSTCKMDITW
FPDQDCENKFGSWTDGQDLVLNSDEGDLSDFTNGEWYLLAMPGGKNTIVYA
CCPEYVDITPTIQRRTLYPFENLIVPCVLISSMALLGFTLPDGSKEKLTGVTL
LSITVFLNLVAESMPTSDAVPLIGTFCNCFMFWASSVLTIVVLNVRHTADITHM
PPMTKSVLOWLPWILEMGRGPKITKTKTLLSNRKELEKERSKSLLANVLDDID
DFRHTISGQTATGSSAFGRPTTVEEHHTAIGCNHKLHLILKELQFITARMKADD
EAEILGDKKFAAMVVDRLIVFTLTITATVTVLLSAPIIVQ"
793. .794
/gene="nAcRalpha-30D"
/note="results in glycine to asparagine substitution;
compared to the sequence deposited in GenBank Accession
Number AE003626"
/replace="aa"
794
variation
/gene="nAcRalpha-30D"
/note="results in glycine to serine substitution; compared
to variant clone"
/replace="a"
846
variation
/gene="nAcRalpha-30D"
/note="results in asparagine to serine substitution;
compared to the sequence deposited in GenBank Accession
Number AE003626 and in variant clone"
/replace="a"
1314
variation
/gene="nAcRalpha-30D"
/note="compared to variant clone"
/replace="a"
1560
variation
/gene="nAcRalpha-30D"
/note="compared to variant clone"

QY 56 LeuGlnGlnIleIleAaspValaspGluIysAsnGlnLeuLeuIleThrAsnIleTrpLeu 75
 DB 565 CTGCAGCAGATCATCGACGTGGATGAAAGAATCAGATTCTGACCAAAATCGTGGTTA 624
 QY 76 SerLeuGluTrpAsnAspTrpAsnLeuArgTrpAsnAspSerGluTrpGlyGlyValIys 95
 DB 625 AATTGGAGTGGAAACGACATACATCTGCGCTGGAATGAACGGAATACGGCGGGTCAAG 684
 QY 96 AspLeuArgIleThrProAsnLysLeuTrpLysProAspValLeuMetTyrAsnSerAla 115
 DB 685 GATCTACGAAATCACGCCCAACAGCTGTGGAAGCCCGAGCTCTCATGTACACAGCGG 744
 QY 116 AspGluGlyPheAspGlyThrTyrGlnThrAsnValValValArgSerGlyGlySerCys 135
 DB 745 GATGAGGATTCGATGCGATCATCACCAACATTTGTGTGTCACCAATACCGCAGTTGT 804
 QY 136 LeuTyrValProProGlyIlePheLysSerThrCysLysMetAspIleAlaTrpPhePro 155
 DB 805 CTGTACGTGCCCCCTGTGTATCTTCAGAGCAGATGCAAGATAGACATCATCGTGGTCCCA 864
 QY 156 PheAspAspGlnHisCysAspMetLysPheGlySerTrpThrTyrAspGlyAsnGlnLeu 175
 DB 865 TTTGATGACCAACATTCGGAATGAAATTCGGTAGTTGGACTTACGATGGAATCAGTTG 924
 QY 176 AspLeuValLeuLysAspGluAlaGlyGlyAspLeuSerAspPheIleThrAsnGlyGlu 195
 DB 925 GATTGGTTTGAATTCGGAAGATGGAGGGGATCTTTCCGATTTCATACAAATATGGCGAG 984
 QY 196 TrpTyrLeuIleGlyMetProGlyLysLysAsnThrIleThrTyrAlaCysCysProGlu 215
 DB 985 TGGTACTGTGTGTCATGCGCGGAAAGAAATACGATAGTACGCGCTCTGCCCGAGAA 1044
 QY 216 ProTyrValAspValThrPheThrIleMetIleArgArgThrLeuTyrTyrPhePhe 235
 DB 1045 CCATATGCGATATCACTTTACTATACAAATTCGTCGCGTACATTATATATTTTTC 1104
 QY 236 AsnLeuIleValProCysValLeuIleSerSerMetAlaLeuLeuGlyPheThrLeuPro 255
 DB 1105 AATTAAATCGTGCCATGTGTCTAATCTCATCGATGCGCTTACTGGGCTTCACATTGCCG 1164
 QY 256 ProAspSerGlyGlyLysLeuThrLeuGlyValThrIleLeuLeuSerLeuThrValPhe 275
 DB 1165 CCGGATTCGGCGGAGAACTGACCTGGAGTTTCAATCTTATGCTCAGAGTGT 1224
 QY 276 LeuAsnLeuValAlaGluThrLeuProGlnValSerAspAlaIleProLeuLeuGlyThr 295
 DB 1225 CTCACCTTGTAGCTGAGACATTCGCCCAAGTATCTGATGCAATCCCTTGTAGGCACC 1284
 QY 296 TyrPheAsnCysIleMetPheMetValAlaSerSerValValLeuThrValValValLeu 315
 DB 1285 TACTTCAATTGCATCATGTTTCATGTCGCCCTCGTGGTGGTGCACAGTAGTGGTCTC 1344
 QY 316 AsnTyrHisHisArgThrAlaAspIleHisGluMetProGlnTrpIleLysSerValPhe 335
 DB 1345 AACTACCAACCATCGACAGCGGACATTCACAGATGCCACCGTGGATCAATCGTGTTC 1404
 QY 336 LeuGlnTrpLeuProTrpIleLeuArgMetSerArgProGlyLysLysIleThrArgLys 355
 DB 1405 CTACAAATGGCTGCCCTGGATCTTGCATGGTGGTGGACCGGTGGCAAGATTACACGCAA 1464
 QY 356 ThrIleMetMetAsnThrArgMetArgGluLeuGluLeuLysGluArgSerSerLysSer 375
 DB 1465 ACAATACATTAAAGCAATCGCATGAAGAGGTGGAGCTGAAGAGCGCTCTCCCAATCC 1524
 QY 376 LeuLeuAlaAsnValLeuAspIleAspAspPheArgHisGlyProProProAsn 395
 DB 1525 CTGTGGCCCAATGCTCGATCGATCGACATCGACGACGACTTCCGGCAC- 1566
 QY 396 SerThrAlaSerThrGlyAsnLeuGlyProGlyCysSerIlePheArgThrAspPheArg 415
 DB 1567 ---ACAATATCTGGCTCCCAACCGCCATTGGCTCGTCG- 1602
 QY 416 ArgSerPheValArgProSerThrMetGluAspValGlyGlyGlyLeuGlySerHisHis 435

DB 1603 GCAGCTTCGGTCGCCCAACAACGGTGGAGGAGCATCACACGCCCATCGTGCATCAC 1662
 QY 436 ArgGluLeuHisLeuIleLeuArgGluLeuGlnPheIleThrAlaArgMetLysLysAla 455
 DB 1663 AAAGATCTTCATCTAATCTCAAGAAATGCAATTTATTACGCGCGGATGCGCAAGCT 1722
 QY 456 AspGluGluAlaGluLeuIleSerAspTrpLysPheAlaMetValValAspArgPhe 475
 DB 1723 GACGACGAAGCGAATGTATCGCGGATTCGGAAGTTCCGCGCAATGGTTGTGATAGATT 1782
 QY 476 CysLeuPheValPheThrLeuPheThrIleIleAlaThrValAlaValLeuLeuSerAla 495
 DB 1783 TGTTTAATTTGTTTTCACGCTCTTCAGATTATTCGAACGGTTACGCTGCTCTCGCT 1842
 QY 496 ProHisIleIleValGln 501
 DB 1843 CCGCACATAATCGTGCAA 1860
 RESULT 9
 AF321448
 LOCUS
 DEFINITION
 Drosophila melanogaster nicotinic acetylcholine receptor dalpha6 subunit variant type IV (nAcRalpha-30D) mRNA, complete cds,
 alternatively spliced.
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Drosophila melanogaster (fruit fly)
 Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 REFERENCE
 AUTHORS
 TITLE
 Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
 Dalpna5, Dalpna6 and Dalpna7, in Drosophila melanogaster Identify a
 New and Highly Conserved Target of Adenosine Deaminase Acting on
 RNA-Mediated A-to-I Pre-mRNA Editing
 Genetics 160 (4), 1519-1533 (2002)
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 Submitted (15-NOV-2000) Human Anatomy and Genetics-Functional
 Genetics Unit, University of Oxford, South Park Road, Oxford OX1
 30X, UK
 FEATURES
 source
 1..2068
 /organism="Drosophila melanogaster"
 /mol_type="mRNA"
 /db_xref="taxon:7227"
 /chromosome="2"
 /map="30D1"
 /dev_stage="embryo"
 1..2068
 /gene="nAcRalpha-30D"
 379..1908
 /gene="nAcRalpha-30D"
 /note="ion channel; neurotransmitter transmembrane
 receptor; alternatively spliced; contains exons 3a, 3b and
 9b; contains two repeats of the functional domain loop D"
 /codon_start=1
 /product="nicotinic acetylcholine receptor Dalpna6 subunit
 variant type IV"
 /protein_id="AA13395.1"
 /db_xref="GI:20152851"
 /translation="MDSPLFASLSLFLVLIIFLAIKSCQCPHEKRLINHLSTVNTL
 ERPVASEPELVGKGLTLQOIIDVDEKNQLTTNAMLNDEKQLLTINLWLSLEWN
 DYNLRWNETEYGVKDLRTITPNKLWPDVLMYNSADEGFDGTYHTNIIVVKSGSLYV
 PPGFKSTCKMDITWFPDQCEMFKSTYDGNQLDLVNSDEGGLSDFITNGEW
 YLLAMPCKNTIVACCPPEPYVDITFTTIQIRRTTYFFNLIVPCVLISMAILGFTL

PPDSGEKLTGLVTILLSLTVFLNVAETLPQVSDAIPILGTYFNCIMFWVASSVLT
VYNTHRTADIPPEPPKTSVFLQWLPILMRGPRKRITKRTILLSNRMKELELKE
RSSKLLANVLIDDDFRHTISGSQTAGSSASFGRTTVEEHHTAIGNKHOLHLIL
KELQFITARMRKADAEELIGDWKFAAMVVDRECLIVFLFTIATVTVLLSAPHIV
Q"

836

variation

/gene="nAcRalpha-30D"
/note="results in arginine to histidine substitution;
compared to the sequence deposited in GenBank Accession
Number AE003626"
/replace="a"

839

variation

/gene="nAcRalpha-30D"
/note="results in serine to asparagine substitution;
compared to the sequence deposited in GenBank Accession
Number AE003626"
/replace="a"

891

variation

/gene="nAcRalpha-30D"
/note="results in methionine to isoleucine substitution;
compared to the sequence deposited in GenBank Accession
Number AE003626"
/replace="a"

1359

variation

/gene="nAcRalpha-30D"
/note="compared to the sequence deposited in GenBank
Accession Number AE003626"
/replace="c"

1605

variation

/gene="nAcRalpha-30D"
/note="compared to the sequence deposited in GenBank
Accession Number AE003626"
/replace="c"

1728

variation

/gene="nAcRalpha-30D"
/note="compared to the sequence deposited in GenBank
Accession Number AE003626"
/replace="c"

ORIGIN

Alignment Scores:

Pred. No.: 2.9e-200 Length: 2068
Score: 2176.50 Matches: 418
Percent Similarity: 86.18% Conservative: 31
Best Local Similarity: 80.23% Mismatches: 37
Query Match: 82.44% Indels: 35
DB: 3 Gaps: 4

US-09-303-232-6 (1-501) x AF321448 (1-2068)

Qy 2 AlaProMetLeuAlaLeuAlaLeu-----LeuAlaLeuPro 15
Db 385 TCCCGCTGCAGCGTCGCTGCTGTTGCTGTTGATCTTTCTGGGATAATTAA 444
Qy 16 ValSerGluGlnGlyProHisGluLysArgLeuLeuAsnAlaLeuAsnTyrAsn 35
Db 445 GAAGCTGTCAAGAGCTCATGAAGAGCGCTGCTGAACCATCTGCTGCCACCTACAAT 504
Qy 36 ThrLeuGluArgProValAlaAsnGluSerGluProLeuGluValArgPheGlyLeuThr 55
Db 505 ACGCTGAGGAGCCCGTGCCCAATGATCGAGCCCTCGAGGTTAAGTTCGAGACTGACG 564
Qy 56 LeuGlnGlnIleLeuAspVal----- 62
Db 565 CTGACGAGATCATCGACGTGGATGAAGAAATCAGATCTCGACCAAAATGCGTGTTA 624
Qy 63 -----AspGluLysAsnGlnLeuLeuIleThrAsnIleTrpLeuSerLeuGluTrpAsn 80
Db 625 AATTGACGAGAGATCAGCTTCTCATACGAATCTTGGCTTTCTGGTGGAGTGGAAAC 684
Qy 81 AspTyrAsnLeuArgTrpAsnAspSerGluTyrGlyGlyValLysAspLeuArgIleThr 100
Db 685 GACTACAAATCTGCGCTGGAATGAACGGAATACGGCGGGTCAAGGATCTACGATACAG 744

Qy 101 ProAsnLysLeuTyrLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAsp 120
Db 745 CCCAACAGACTGTGAAGCCCGACGTGCTCATGTACACAGCGCGATGAGGATTCGAT 804
Qy 121 GlyThrTyrGlnThrAsnValValValArgSerGlyGlySerCysLeuTyrValProPro 140
Db 805 GGCACGTATCACACCAACATTTGTCGTCAAACGTAGCGGAGTGTCTGTACGTGCCCT 864
Qy 141 GlyIlePheLysSerThrCysLysMetAspIleAlaTrpPheProPheAspGlnHis 160
Db 865 GGTATCTTCAAGAGACACATGCAAGATGGACATCACTGCTGCCATTTCCATGACCAACAT 924
Qy 161 CysAspMetLysPheGlySerTrpTyrAspGlyAsnGlnLeuAspLeuValLeuLys 180
Db 925 TCCGAATTAATAATTCGGTAGTTGGACTTACGATGGAATCAGTTGGATTTGTTGAAT 984
Qy 181 AspGluAlaGlyGlyAspLeuSerAspPheIleThrAsnGlyGluTyrTrpLeuIleGly 200
Db 985 TCCGAAGATGGAGGGGATCTTTCCGATTTTATAACAATGGCGAGTGGTACTTGTGTC 1044
Qy 201 MetProGlyLysLysAsnThrIleThrTyrAlaCysCysProGluProTyrValAspVal 220
Db 1045 ATGCGGGAAGAAATACGATATGCTACGCTGCTGCCAGAACCATATGTCGATATC 1104
Qy 221 ThrPheThrIleMetIleArgArgThrLeuTyrTyrPhePheAsnLeuIleValPro 240
Db 1105 ACCTTTACTATACAAATTCGTCGCCGATATATATATTTTCAATTAATGTCGCA 1164
Qy 241 CysValLeuIleSerSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGlu 260
Db 1165 TGTGTGCTAATCTCATCGATGGCCTACTGGGCTTCACATTCGCGCGGATTCGGCGAG 1224
Qy 261 LysLeuThrLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnLeuValAla 280
Db 1225 AAATGAGCGTGGGAGTTACAAATTTCTTATCGCTCACAGTGTCTTCAACCTGTAGCT 1284
Qy 281 GluThrLeuProGlnValSerAspAlaIleProLeuLeuGlyThrTyrPheAsnCysIle 300
Db 1285 GAGACATTCGCCCAAGTATCTGATGCAATCCCTTTGTAGGACCTACTTCAATGTCATC 1344
Qy 301 MetPheMetValAlaSerSerValLeuThrValValLeuAsnTyrHisArg 320
Db 1345 ATGTTTCATGTGCGCATCGTGGTGTGTCGACAGTAGTGGTGTCTCACTACCACTGCG 1404
Qy 321 ThrAlaAspIleHisGluMetProGlnTrpIleLysSerValPheLeuGlnTrpLeuPro 340
Db 1405 ACAGCGGACATTCAGAGATGCCACCGTGGATCAAGTCCGTTTTCTTACAAATGCTGCC 1464
Qy 341 TrpIleLeuArgMetSerArgProGlyLysLysIleThrArgLysThrIleMetMetAsn 360
Db 1465 TGGATCTTGGATGGTTCGACCGTTCGCAAGATTACGCAAAACAATCTATTAAAGC 1524
Qy 361 ThrArgMetArgGluLeuGluLeuLysGluArgSerSerLysSerLeuLeuAlaAsnVal 380
Db 1525 AATCGCATGAAGAGCTGGAGCTAAAGAGCGCTCTCTCCAAATCCCTGCTGGCCAAATGC 1584
Qy 381 LeuAspIleAspAspPheArgHisGlyProProProProAsnSerThrAlaSerThr 400
Db 1585 CTGACATCAGCAGCACTTTCCGCAC-----ACAATATCTGGC 1623
Qy 401 GlyAsnLeuGlyProGlyCysSerIlePheArgThrAspPheArgArgSerPheValArg 420
Db 1624 TCCCAAAACCCCATTTGGCTGTCG-----GCCACTTCGGTCGG 1662
Qy 421 ProSerThrMetGluAspValGlyGlyLeuGlySerHisHisArgGluLeuHisLeu 440
Db 1663 CCCACACGCTGAGGAGCATCACACGGCCATCGGTGCAATCACAAGATCTTCATCTA 1722
Qy 441 IleLeuArgGluLeuGlnPheIleThrAlaArgMetLysLysAlaAspGluAlaGlu 460
Db 1723 ATTCTTAAGATTGCAATTTATTACGGCGCGATGCGAAAGCTGACGACGAGCGGAA 1782

```

Qy 461 LeuileSerAspTrpIysPheAlaAalaMetValValAspArgPheCysLeuPheValPhe 480
Db 1783 TTGATCGCGGATTGGGAAGTTGCGGCAATGGTTGTGGATAGATTGTTTAATTGTTTC 1842
Qy 481 ThrLeuPheThrIleAlaThrValAlaValLeuLeuSerAlaProHisIleIleVal 500
Db 1843 ACGCTCTTCAGATTATTGCAACGGTTACGGTGTCTCTCCGCTCGCACATAAATCGTG 1902
Qy 501 Gln 501
Db 1903 CAA 1905

RESULT 10
AF321449
LOCUS
DEFINITION
  2110 bp mRNA linear INV 29-APR-2002
  Drosophila melanogaster nicotinic acetylcholine receptor Dalp6a6
  subunit variant type V (nAcRalpha-30D) mRNA, complete cds,
  alternatively spliced.
ACCESSION
  AF321449
VERSION
  AF321449.1 GI:20152852
KEYWORDS
  Drosophila melanogaster (fruit fly)
ORGANISM
  Drosophila melanogaster
  Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
  Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
  Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
  1 (bases 1 to 2110)
  Grauso, M., Reenan, R.A., Culetto, E. and Sattelle, D.B.
  Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
  Dalp6a5, Dalp6a6 and Dalp6a7, in Drosophila melanogaster Identify a
  New and Highly Conserved Target of Adenosine Deaminase Acting on
  RNA-Mediated A-to-I Pre-mRNA Editing
  Genetics 160 (4), 1519-1533 (2002)
JOURNAL
  MEDLINE
  21969411
PUBMED
  11973307
REFERENCE
  2 (bases 1 to 2110)
  Grauso, M. and Sattelle, D.B.
  Direct Submission
  Submitted (15-NOV-2000) Human Anatomy and Genetics-Functional
  Genetics Unit, University of Oxford, South Park Road, Oxford OX1
  3QX, UK
FEATURES
  source
    1..2110
    /organism="Drosophila melanogaster"
    /mol_type="mRNA"
    /db_xref="taxon:7227"
    /chromosome="2"
    /map="30D1"
    /dev_stage="embryo"
  gene
    1..2110
    /gene="nAcRalpha-30D"
    379..1950
    /gene="nAcRalpha-30D"
    /note="ion channel; neurotransmitter transmembrane
    receptor; alternatively spliced; contains exons 3a, 8a and
    8b; contains two repeats of the transmembrane TM2 region"
    /codon_start=1
    /product="nicotinic acetylcholine receptor Dalp6a6 subunit
    variant type V"
    /protein_id="AAAI3396.1"
    /db_xref="GI:20152853"
    /translation="MDSELPASLSFLVLLIFLAILKESQCPHEKRLNHLSTYNTIL
    ERPVANESEPELVKFLTLQIIDVEKNQILITNAMLNLEWINDYLNKWNRETYGGVK
    DLRTIPNKLKPDVLMYNSADEGFGYHTNVIVKHNGSCLYVPPGIFKSTCKMDITW
    PFDQHCMEKFGSYTDGNQDLVNSDEGDLSDFITNGEWYLLAMPKKNITIVA
    CCPYVDITITQIRRTLYFFNLIIVPCVLISSMALLGFTLPPDSGEKLTGLVTIL
    LSLTFVLNLVAESPEPTSDAVPLGVTLISLTVFLNLVAETLPOVSDAIELLGTVEN
    CIMFWASSVVLTVVLNHHRTADIEHPMPWIKSVFLQWLPILRMGRPGRKTRKT
    ILLSNRMKELEKRSKLLANVLIDDDPFRTISGTAIGSSASGRPTVEEH
    TAIGNKDLHLILKELOFITARMRKADDEAELIGDWKFAAMVVDVRFCLIVFTLFTII
    ATVTLLISAPHIIQV"
    846
    /gene="nAcRalpha-30D"
  variation

```

```

/note="results in methionine to isoleucine substitution;
compared to the sequence deposited in GenBank Accession
Number AE003626"
/replace="a"
1401
/gene="nAcRalpha-30D"
/note="compared to the sequence deposited in GenBank
Accession Number AE003626"
/replace="c"
1647
/gene="nAcRalpha-30D"
/note="compared to the sequence deposited in GenBank
Accession Number AE003626"
/replace="c"
1770
/gene="nAcRalpha-30D"
/note="compared to the sequence deposited in GenBank
Accession Number AE003626"
/replace="c"
variation
variation
variation
ORIGIN
Alignment Scores:
Pred. No.: 2,55e-198 Length: 2110
Score: 2156.50 Matches: 415
Percent Similarity: 83.55% Conservative: 32
Best Local Similarity: 77.57% Mismatches: 39
Query Match: 81.69% Indels: 49
DB: 3 Gaps: 4
US-09-303-232-6 (1-501) x AF321449 (1-2110)
Qy 2 AlaProMetLeuAlaAlaLeuAlaLeu-----LeuAlaLeuLeuPro 15
Db 385 TCCCGCTGCCAGCTGCTGCTGCTTTGCTCTGTGATCTTCTGGCGATTAATAA 444
Qy 16 ValSerGluGlnGlyProHisGluLysArgLeuLeuAsnAlaLeuLeuAsnTyrAsn 35
Db 445 GAAAGCTGTCAAGGACCTCATGAAAGCGCTGCTGAACCATCTGCTGCCACCTACAAT 504
Qy 36 ThrLeuGluArgProValAlaAsnGluSerGluProLeuGluValArgPheGlyLeuThr 55
Db 505 ACGCTGGAGCGACCCGTGGCAATGAATCGAGCCCTGGAGGTTAAGTTTCGGACTGACG 564
Qy 56 LeuGlnGlnIleIleAspValAspGluLysAsnGlnLeuLeuIleThrAsnIleTrpLeu 75
Db 565 CTGCAGCAGATCATCGACGTGGATGAAAGAAATCAGATTCTGACCAATATGCGTGGTTA 624
Qy 76 SerLeuGluTrpAsnAspTyrAsnLeuArgTrpAsnAspSerGluTyrGlyValIys 95
Db 625 AATTGGAGTGGAAACGACTACAAATCTGGCTGGAAATGAAACGGAATACGGCGGGTCAAG 684
Qy 96 AspLeuArgIleThrProAsnLysLeuTrpLysProAspValLeuLeuMetTyrAsnSerAla 115
Db 685 GATCTAGAAATCAAGCCCAACAGCTGTGGAAGCCGACGCTCATGTACAAACAGCGCG 744
Qy 116 AspGluGlyPheAspGlyThrTyrGlnThrAsnValValArgSerGlySerCys 135
Db 745 GATGAGGATTTCGATGGCAGCATATCACCCNACAAATTGTGTGTCAAACATAACGGCAGTTGT 804
Qy 136 LeuTyrValProProGlyIlePheLysSerThrCysLysMetAspIleAlaTrpPhePro 155
Db 805 CTGTACGTGCCCCCTGGTATCTTCAAGAGCACATGCAAGATGGACATCACGTGTTCCCA 864
Qy 156 PheAspAspGlnHisCysAspMetLysPheGlySerTrpThrTyrAspGlyAsnGlnLeu 175
Db 865 TTTGATGACCACAACTTCGAAATGAAATTCGGTAGTTGGACTTACGATGGAATACAGTTG 924
Qy 176 AspLeuValLeuLysAspGluAlaGlyAspLeuSerAspPheIleThrAsnGlyLeu 195
Db 925 GATTGTGTTTGAATTCGAGAGATGGAGGGGATCTTCCCGATTTTCATACCAATATGGCGAG 984
Qy 196 TrpTyrLeuIleGlyMetProGlyLysLysAsnThrIleThrTyrAlaCysCysProGlu 215

```


Db	985	TGGTACTTCTGCTGCGGAAAGAAATACGATAGTCTAGCGCTGCTGCCAGAA	1044
Qy	216	ProTyrValAspValThrPheThrIleMetIleArgArgArgThrLeuTyrTyrPhePhe	235
Db	1045	CCATATGCGATATACCTTTACTATACAAATCGTCGCGGACATATATATTTTC	1104
Qy	236	AsnLeuIleValProCysValLeuIleSerSerMetAlaLeuLeuGlyPheThrLeuPro	255
Db	1105	AATTTAATCGTCGCGATGTCATCTCATCGATGCGCTACTGGCTTCACATGCGG	1164
Qy	256	ProAspSerGlyGluLeuThrLeu	264
Db	1165	CGGATTCGGGCGAAGAACTGACGCTGCGCGTAACACTACTACTATCATTAACAGTATT	1224
Qy	265	-----GlyVal	266
Db	1225	CTAAACCTTGTGCGCGAGTCCATGCGGACAGCTCGGATGCTGTTCTTATAGAGTT	1284
Qy	267	ThrIleLeuLeuSerLeuThrValPheLeuAsnLeuValAlaGluThrLeuProGlnVal	286
Db	1285	ACAATTCTTCTATCGTCCACAGTGTCTCAACCTTGTAGCTGAGACATTGCCCAAGTA	1344
Qy	287	SerAspAlaIleProLeuLeuGlyThrTyrPheAsnCysIleMetPheMetValAlaSer	306
Db	1345	TCTGATGCAATCCCTGTTAGGCACCTACTTCAATTCATCATCTGTCATGTCGATCG	1404
Qy	307	SerValValLeuThrValValLeuAsnTyrHisArgThrAlaAspIleHisGlu	326
Db	1405	TCGGTGTGCTGACAGTAGTGTCTCACTACCACTCCACAGCGGACATTCCAGG	1464
Qy	327	MetProGlnTfPileLysSerValPheLeuGlnTfPileLeuArgMetSer	346
Db	1465	ATGCCACCGTGGATCAAGTCGCTTTCTCAATGGCTGCTCGATCTCGAATGGT	1524
Qy	347	ArgProGlyLysLysIleThrArgLysThrIleMetMetAsnThrArgMetArgGluLeu	366
Db	1525	CGACCCGTCGACAGATTACAGCAAAACATATATTAAAGCAATCGCATGAAGAGCTG	1584
Qy	367	GluLeuLysGluArgSerSerLysLeuAlaAsnValLeuAspIleAspAsp	386
Db	1585	GAGCTAAGGAGCGCTCTCCAAATCCCTGCTGGCCATGCTCTCGACATCGACGAC	1644
Qy	387	PheArgHisGlyProProProProAsnSerThrAlaSerThrGlyAsnLeuGlyProGly	406
Db	1645	TTTCGGCAC-----ACAATATCTGGCTCCAAACCGCCATTGGC	1683
Qy	407	CysSerIlePheArgThrAspPheArgSerPheValArgProSerThrMetGluAsp	426
Db	1684	TCGTCTG-----GCCAGCTTCGTCGCGCCACCAACCGTGGAGAG	1722
Qy	427	ValGlyGlyLeuGlySerHisArgGluLeuHisLeuIleLeuArgGluLeuGln	446
Db	1723	CATCACCGGCGCATCGCTGCAATCACAAAGATCTTCATCTAATCTTAAGAATTGCA	1782
Qy	447	PheIleThrAlaArgMetLysLysAlaAspGluAlaGluLeuIleSerAspTrpLys	466
Db	1783	TTTATTACGGCGCGGATGCGCAAGCTGACGACGAGCGGAATTGATCGCGGATTGGAAG	1842
Qy	467	PheAlaAlaMetValValAspArgPheCysLeuPheValPheThrLeuPheThrIle	486
Db	1843	TTCCGCGCAATGTTGTGGATAGATTGTTTAATTGTTTTCAGCTCTTCACGATTATT	1902
Qy	487	AlaThrValAlaValLeuSerAlaProHisIleIleValGln	501
Db	1903	GCAACGGTTACGGTCTGCTCTCCGCTCCGACATAATCGTGCAA	1947
RESULT 11			
AF143846			
LOCUS	AF143846	3629 bp	mRNA
DEFINITION	Heliothis virescens putative nicotinic acetylcholine receptor alpha		
ACCSSION	AF143846		
VERSION	AF143846.1	GI:4895004	

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL